GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:36:19 ; Search time 82 Seconds
(without alignments)
1118.183 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2370	99.7	445	10	Q8W5B2	Q8w5b2 arabidopsis
2	2285	96.1	440	10	Q94CF5	Q94cf5 arabidopsis
3	923	38.8	631	10	Q93V59	Q93v59 arabidopsis
4	615	25.9	223	10	Q9LU50	Q9lu50 arabidopsis
5	552	23.2	107	10	Q8W5B1	Q8w5b1 arabidopsis
6	534.5	22.5	632	10	Q9ZQP0	Q9zqp0 arabidopsis
7	510.5	21.5	692	10	Q9ZNT9	Q9znt9 arabidopsis
8	451.5	19.0	186	10	Q94IE6	Q94ie6 arabidopsis
9	423	17.8	851	10	O23524	O23524 arabidopsis
10	322.5	13.6	623	10	O23521	O23521 arabidopsis
11	266	11.2	295	10	O23526	O23526 arabidopsis
12	221	9.3	739	4	Q96BD9	Q96bd9 homo sapien
13	221	9.3	803	4	Q15022	Q15022 homo sapien
14	219	9.2	855	5	Q8T9D8	Q8t9d8 drosophila
15	219	9.2	900	5	Q9NJG9	Q9njg9 drosophila
16	219	9.2	955	5	Q9VW55	Q9vw55 drosophila
17	180	7.6	388	10	Q9LU51	Q9lu51 arabidopsis
18	168	7.1	1231	5	O97140	O97140 dictyosteli
19	156.5	6.6	369	5	Q8T1D9	Q8t1d9 dictyosteli
20	153.5	6.5	1309	5	Q8T2H9	Q8t2h9 dictyosteli
21	147.5	6.2	4550	5	O77336	O77336 plasmodium
22	143	6.0	530	5	Q8T2A0	Q8t2a0 dictyosteli
23	141.5	6.0	898	5	Q8T1G7	Q8t1g7 dictyosteli
24	140.5	5.9	548	5	Q8T1S2	Q8t1s2 dictyosteli
25	140.5	5.9	1204	5	Q8T134	Q8t134 dictyosteli
26	140.5	5.9	1711	5	O77322	O77322 plasmodium
27	139.5	5.9	1844	5	O97287	O97287 plasmodium
28	138	5.8	682	5	Q9GPR8	Q9gpr8 dictyosteli
29	136	5.7	855	5	Q8SSV6	Q8ssv6 dictyosteli
30	136	5.7	1318	5	Q95PH4	Q95ph4 dictyosteli
31	136	5.7	1887	5	Q8SSS8	Q8sss8 dictyosteli
32	135	5.7	164	10	Q9XFE7	Q9xfe7 oryza sativ
33	135	5.7	1208	5	O97101	O97101 dictyosteli
34	134	5.6	261	5	Q8T868	Q8t868 dictyosteli
35	134	5.6	1529	5	Q9GQC2	Q9gqc2 dictyosteli
36	133.5	5.6	1364	5	Q8T223	Q8t223 dictyosteli
37	133.5	5.6	3848	5	O76737	O76737 dictyosteli
38	133	5.6	1903	5	Q9U0N7	Q9u0n7 plasmodium
39	131.5	5.5	1245	5	O96195	O96195 plasmodium
40	131.5	5.5	2472	5	Q8T2M5	Q8t2m5 dictyosteli
41	131	5.5	340	5	Q8T212	Q8t212 dictyosteli
42	131	5.5	1407	5	Q8T1T5	Q8t1t5 dictyosteli
43	131	5.5	2994	5	Q95ZG5	Q95zg5 dictyosteli
44	130.5	5.5	699	5	O77318	O77318 plasmodium
45	130.5	5.5	721	5	Q8T218	Q8t218 dictyosteli

ALIGNMENTS

RESULT 1
Q8W5B2
ID Q8W5B2 PRELIMINARY; PRT; 445 AA.
AC Q8W5B2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Vernalization 2 protein.
GN VRN2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21575875; PubMed=11719192;
RA Gendall A.R., Levy Y.Y., Wilson A., Dean C.;
RT "The VERNALIZATION2 (VRN2) gene mediates the epigenetic regulation of
RT vernalization in Arabidopsis.";
RL Cell 107:525-535(2001).
DR EMBL; AF284500; AAL32135.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 445 AA; 51177 MW; 30A3391CE98D877C CRC64;

Query Match 99.7%; Score 2370; DB 10; Length 445;
Best Local Similarity 99.8%; Pred. No. 5.2e-188;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180
```

```
Db        121 SEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240

Db        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360

Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420

Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420

Qy        421 DHPNDINNKNNVDNKDNNSRDKVIK 445

Db        421 DHPNDINNKNNVDNKDNNSRDKVIK 445


RESULT 2
Q94CF5
ID   Q94CF5       PRELIMINARY;      PRT;   440 AA.
AC   Q94CF5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 50.6 kDa protein.
GN   DL4450W.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA   Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA   Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA   Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA   Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "Full Length cDNA of gene dl4450w.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA   Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA   Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA   Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA   Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA   Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RT   "Full Length cDNA of gene dl4450w.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY034902; AAK59409.1; -.
DR   EMBL; AY063047; AAL34221.1; -.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW   DNA-binding; Hypothetical protein; Zinc-finger.
SQ   SEQUENCE   440 AA;  50623 MW;  0E0AB2C55517BE6F CRC64;

  Query Match              96.1%;  Score 2285;  DB 10;  Length 440;
  Best Local Similarity    96.8%;  Pred. No. 5.6e-181;
  Matches  428;  Conservative    3;  Mismatches    9;  Indels    2;  Gaps    1;
```

```
Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Db         61 KSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Db        121 LEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240

Db        181 SPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360

Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEVFSRFYEKELHCYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420

Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNTS--VTNNNNNSVDHPSDSNTNNNNIV 418

Qy        421 DHPNDINNKNNVDNKDNNSRDK 442

Db        419 DHPNDIKNKNNVDNKDNNSRDK 440


RESULT 3
Q93V59
ID   Q93V59       PRELIMINARY;      PRT;   631 AA.
AC   Q93V59;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Embryonic flower 2.
GN   EMF2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshida N., Yanai Y.;
RT   "Arabidopsis, embryonic flower 2 gene.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yoshida N.;
RT   "embryonic flower 2.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB053262; BAB58957.1; -.
DR   EMBL; AB053171; BAB58956.1; -.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   631 AA;  71680 MW;  F36FE92D0F62E610 CRC64;

  Query Match              38.8%;  Score 923;  DB 10;  Length 631;
  Best Local Similarity    34.8%;  Pred. No. 5.4e-68;
  Matches  216;  Conservative   52;  Mismatches  108;  Indels  244;  Gaps    9;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Db         21 MCHEDSRLRISEEEEIAAEESLAAYCKPVELYNIIQRRAIRNPLFLQRCLHYKIEAKHKR 80
```

```
Qy          61 ------------------------------------------------------------ 60

Db          81 RIQMTVFLSGAIDAGVQTQKLFPLYILLARLVSPKPVAEYSAVYRFSRACILTGGLGVDG 140

Qy          61 ------------------------------------------------------------ 60

Db         141 VSQAQANFLLPDMNRLALEAKSGSLAILFISFAGAQNSQFGIDSGKIHSGNIGGHCLWSK 200

Qy          61 -------------------------------------KSRS------------------ 64
                                                    || |
Db         201 IPLQSLYASWQKSPNMDLGQRVDTVSLVEMQPCFIKLKSMSEEKCVSIQVPSNPLTSSSP 260

Qy          65 ---------------------------------------TGMVVFNYKDCNNTLQKTEV 84
                                                      || ||||:  || ||||||
Db         261 QQVQVTISAEEVGSTEKSPYSSFSYNDISSSSLLQIIRLRTGNVVFNYRYYNNKLQKTEV 320

Qy          85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSD 144
                || ||||| : | |||||::|| |:|||  ||| : ||:| ||||:|  : | :    |
Db         321 TEDFSCPFCLVKCASFKGLRYHLPSTHDLLNFEFWVTEEFQAVNVSLKTETMISKVNEDD 380

Qy         145 -DDKFEPFSLCSKP-RKRRQRGGRNNTRRLKVCFLPLDSPSLTNG------TENGITLLN 196
                | | : |   ||  |:|||:    ::|:          | |  |      |::  :: :
Db         381 VDPKQQTFFFSSKKFRRRRQKSQVRSSRQ---------GPHLGLGCEVLDKTDDAHSVRS 431

Qy         197 DGNR---GLGYPEATELAGQFEMTSNIP----PAIAHSSLDAGAKVILTSEAVVPATKTR 249
               : :|   |  |    |  | |   :|    ||   |  |      :    ::   |||
Db         432 EKSRIPPGKHY----ERIGGAESGQRVPPGTSPADVQSCGDPDYVQSIAGSTMLQFAKTR 487

Qy         250 KLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDF 309
               |:| |||: |:  |||||||:|||| ||||||||:||||||||||||||||||:|||||
Db         488 KISIERSDLRNRSLLQKRQFFHSHRAQPMALEQVLSDRDSEDEVDDDVADFEDRRMLDDF 547

Qy         310 VDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIK 369
               ||| ||||| ||:||||||||||:||||| |||||||| :   : |   | ||||:|::|
Db         548 VDVTKDEKQMMHMWNSFVRKQRVLADGHIPWACEAFSRLHGPIMVRTPHLIWCWRVFMVK 607

Qy         370 LWNHGLVDSATINNCNTILE 389
               ||||||:|: |:||||| ||
Db         608 LWNHGLLDARTMNNCNTFLE 627


RESULT 4
Q9LU50
ID   Q9LU50      PRELIMINARY;      PRT;   223 AA.
AC   Q9LU50;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Similarity to fertilization-independent seed 2 protein.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=20181125; PubMed=10718197;
RA   Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT   features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:31-63(2000).
DR   EMBL; AB023044; BAA97387.1; -.
SQ   SEQUENCE   223 AA;  26064 MW;  0E68687EA4BA9B60 CRC64;

  Query Match             25.9%;  Score 615;  DB 10;  Length 223;
  Best Local Similarity   62.7%;  Pred. No. 4.7e-43;
  Matches  116;  Conservative   20;  Mismatches   45;  Indels    4;  Gaps    1;

Qy         209 ELAGQFEMTSNIP----PAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLL 264
               |  |  |    :|    ||   |  |      :  . ::   ||||:| |||: |:  ||
Db          35 ERIGGAESGQRVPPGTSPADVQSCGDPDYVQSIAGSTMLQFAKTRKISIERSDLRNRSLL 94

Qy         265 QKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWN 324
               |||||:|||| ||||||||:|||||||||||||||||||:|||||||| ||||| ||:||
Db          95 QKRQFFHSHRAQPMALEQVLSDRDSEDEVDDDVADFEDRRMLDDFVDVTKDEKQMMHMWN 154

Qy         325 SFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNC 384
               |||||||||:|||| ||||||| :   : |   | ||||:|::|||||||:|: |:|||
Db         155 SFVRKQRVLADGHIPWACEAFSRLHGPIMVRTPHLIWCWRVFMVKLWNHGLLDARTMNNC 214

Qy         385 NTILE 389
               || ||
Db         215 NTFLE 219


RESULT 5
Q8W5B1
ID   Q8W5B1      PRELIMINARY;      PRT;   107 AA.
AC   Q8W5B1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Vernalization 2 protein.
GN   VRN2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21575875; PubMed=11719192;
RA   Gendall A.R., Levy Y.Y., Wilson A., Dean C.;
RT   "The VERNALIZATION2 (VRN2) gene mediates the epigenetic regulation of
RT   vernalization in Arabidopsis.";
RL   Cell 107:525-535(2001).
DR   EMBL; AF284501; AAL32136.1; -.
SQ   SEQUENCE   107 AA;  12393 MW;  2B3321C2BB6380F4 CRC64;

  Query Match             23.2%;  Score 552;  DB 10;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 3e-38;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy          61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101
               |||||||||||||||||||||||||||||||||||||||||
Db          61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101


RESULT 6
Q9ZQP0
ID   Q9ZQP0      PRELIMINARY;      PRT;   632 AA.
AC   Q9ZQP0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Fertilization-independent seed 2 protein.
GN   AT2G35670.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA   Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA   Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA   Fraser C.M., Venter J.C.;
```

```
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC006068; AAD15448.2; -.
SQ   SEQUENCE   632 AA;  71130 MW;  DDFDE33DC6620AE8 CRC64;

  Query Match           22.5%;  Score 534.5;  DB 10;  Length 632;
  Best Local Similarity   23.8%;  Pred. No. 8.1e-36;
  Matches  151;  Conservative   61;  Mismatches  113;  Indels  309;  Gaps   10;

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVSVKLNSFIF 138
              | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|||| ::  |
Db          4 KAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDVSVKSDTIKF 63

Qy        139 E--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195
                ::   : :  | : ||| | :|||| || ::| |  : ||   |  |||| | :
Db         64 GVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDDLPRGTENDSTHV 123

Qy        196 NDGN---------------------------------------------------------- 199
              || |
Db        124 NDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAHSSAEKNEST 183

Qy        200 ----------------------------------------------------RGLGYPEATEL 210
                                                                   : :|  |
Db        184 HVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDEDVSFPPRTRS 243

Qy        211 AGQFEMTSNI-----------------------------------PPAIAHS-------- 227
              : :   ||:|                                  |||| |
Db        244 SKE---TSDILTTTQPAIVEPSEPKVRRVSRRKQLYAKRYKARETQPAIAESSEPKVLHV 300

Qy        228 ------------SLDAGAKVILTSEAVV------------------------------- 243
                         ||:  : :: |::  :
Db        301 NDENVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSKKNKSTR 360

Qy        244 ------------------------------------------------------------ 243

Db        361 KNVDNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSSKKNKST 420

Qy        244 ----------PATKTRK----------------------------LSAERSEARSHLLL 264
                       | |:: |                            | ||| ||:    |
Db        421 RKNDDNIPSPPKTRSSKKTSNILATTQPAKAEPSEPKVTRVSRRKELHAERCEAKRLERL 480

Qy        265 QKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWN 324
              : |||||  :||:  |||||: |||:| ||   |  :|  |:  | |:|:||::|:|:|||
Db        481 KGRQFYHSQTMQPITFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKRYMYLWN 540

Qy        325 SFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNC 384
               |||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: : | : |
Db        541 IFVRKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICAKTFHKC 600

Qy        385 NTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
               |||          :|:: :   | |  | ||
Db        601 TTIL----------LSNSDEAGQFTSGSAANANN 624


RESULT 7
Q9ZNT9
ID   Q9ZNT9      PRELIMINARY;      PRT;   692 AA.
AC   Q9ZNT9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   FERTILIZATION-independent seed 2 protein.
GN   FIS2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99093530; PubMed=9874812;
RA   Luo M., Bilodeau P., Koltunow A., Dennis E.S., Peacock W.J.,
RA   Chaudhury A.;
RT   "Genes controlling fertilization-independent seed development in
RT   Arabidopsis thaliana.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:296-301(1999).
DR   EMBL; AF096096; AAD09105.1; -.
DR   EMBL; AF096095; AAD09104.1; -.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
FT   VARIANT     637     637       W -> *.
SQ   SEQUENCE   692 AA;  77666 MW;  2A3F61B961676497 CRC64;

  Query Match           21.5%;  Score 510.5;  DB 10;  Length 692;
  Best Local Similarity   22.3%;  Pred. No. 8.8e-34;
  Matches  154;  Conservative   53;  Mismatches  121;  Indels  363;  Gaps    8;

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVSVKLNSFIF 138
              | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|||| ::  |
Db          4 KAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDVSVKSDTIKF 63

Qy        139 E--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195
                ::   : :  | : ||| | :|||| || ::| |  : ||   |  |||| | :
Db         64 GVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDDLPRGTENDSTHV 123

Qy        196 NDGN---------------------------------------------------------- 199
              || |
Db        124 NDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAHSSAEKNEST 183

Qy        200 ------------------------------------------------------------ 199

Db        184 HVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDEDVSFPPRTRS 243

Qy        200 ------------------------RG-------------------------------- 201
                                       ||
Db        244 SKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSEPKVLHVNDE 303

Qy        202 --LGYPEATELAGQFEMTSNIPPAIAHSS------------------------------ 228
                  |||  |    :: :   |||| ||
Db        304 NVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSKKNKSTRKNV 363

Qy        229 ------------------------------------------------------------ 228

Db        364 DNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSSKKNKSTRKN 423

Qy        229 -------------------LDAGAKVILTSEAVVP------------------------ 244
                              |     |  || ||
Db        424 DDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSSKKNKSTHKK 483

Qy        245 -----------------------------------ATKTRKLSAERSEARSHLLLQKR 267
                                                  :: ::| ||| ||:    |: |
Db        484 DDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEAKRLERLKGR 543

Qy        268 QFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
              |||||  :|||  |||||: |||:| ||   |  :|  |:  | |:|:||::|:|:||| ||
Db        544 QFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKRYMYLWNIFV 603

Qy        328 RKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTI 387
              |||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: : | : | ||
Db        604 RKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICAKTFHKCTTI 663

Qy        388 LENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
              |           :|:: :    | |  | ||
Db        664 L----------LSNSDEAGQFTSGSAANANN 684


RESULT 8
Q94IE6
ID   Q94IE6      PRELIMINARY;      PRT;   186 AA.
AC   Q94IE6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Embryonic flower 2 like 1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yoshida N., Yanai Y.;
RT   "Arabidopsis, embryonic flower 2 like 2 gene.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB053265; BAB58960.1; -.
SQ   SEQUENCE   186 AA;  21754 MW;  D629EB63E9381853 CRC64;

  Query Match             19.0%;  Score 451.5;  DB 10;  Length 186;
  Best Local Similarity   50.8%;  Pred. No. 1.3e-29;
  Matches   99;  Conservative   32;  Mismatches   43;  Indels   21;  Gaps    7;

Qy        237 LTSEAVVPATKTRKLSAERSEARSHLL-LQKRQFYHSHRVQPMALEQVMSDRDSEDEVD- 294
              ||:|| ||        |:||:| || | | |||||||   ||::||||||||||::||
Db          6 LTTEAKVP--------AKRSKATSHYLPLHKRQFYHSRTGQPLSLEQVMSDRDSENDVDK 57

Qy        295 -DDVADFEDRQMLDDFVDVNK-DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKE 352
               || |  |: |||:  :| |:   ::|: ||||||::||::|| || |||||||| : :|
Db         58 NDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAHIPWACEAFSRLHLQE 117

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDS 412
              |    ||  ||| |:|| |::||:|  |:| ||||: :  | | |  . |||:
Db        118 LRSNLSLDLCWRQFMIKQWDYGLLDRVTMNKCNTIIYH--NISTTNDDINNNNT-----R 170

Qy        413 NTNNNNIVDHPNDIN 427
              |:| ::||  :|||
Db        171 TTDNMDVVD--DDIN 183


RESULT 9
O23524
ID   O23524      PRELIMINARY;      PRT;   851 AA.
AC   O23524; O23525;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 96.8 kDa protein.
GN   AT4G16840.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA   Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA   Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA   Jones J., Palme K., Ansorge W., Delseny M., Bancroft I., Mewes H.W.,
RA   Schueller C., Chalwatzis N.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97342; CAB10457.1; -.
DR   EMBL; AL161545; CAB80955.1; -.
DR   InterPro; IPR002885; PPR.
DR   Pfam; PF01535; PPR; 9.
DR   TIGRFAMs; TIGR00756; PPR; 10.
KW   Hypothetical protein.
SQ   SEQUENCE   851 AA;  96849 MW;  9BF5E46E5FFEB71C CRC64;

  Query Match             17.8%;  Score 423;  DB 10;  Length 851;
  Best Local Similarity   39.6%;  Pred. No. 2e-26;
  Matches  101;  Conservative    2;  Mismatches    0;  Indels  152;  Gaps    2;

Qy         43 PSFLPRCLNYKIGAKRKRK-----------------------------------SRS 64
              ||||||||||||||||||                                   |||
Db        701 PSFLPRCLNYKIGAKRKRKYAFLLECSCHSDMLFILLLIWKLMNYLSLLSRYGHNEWSRS 760

Qy         65 TGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEY 124
              ||||||||||||||||:||
Db        761 TGMVVFNYKDCNNTLQRTE---------------------------------------- 779

Qy        125 QTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSL 184

Db        780 ------------------------------------------------------------ 779

Qy        185 TNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVP 244
                                  ||||||||||||||||||||||||||||||||||||||:|||||
Db        780 ------------GNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTEAVVP 826

Qy        245 ATKTRKLSAERSEAR 259
              |||||||||||||||
Db        827 ATKTRKLSAERSEAR 841


RESULT 10
O23521
ID   O23521      PRELIMINARY;      PRT;   623 AA.
AC   O23521;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   Hypothetical 70.1 kDa protein.
GN   AT4G16810.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA   Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA   Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA   Jones J., Palme K., Ansorge W., Delseny M., Bancroft I., Mewes H.W.,
RA   Schueller C., Chalwatzis N.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97342; CAB10454.1; -.
DR   EMBL; AL161545; CAB80952.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   623 AA;  70062 MW;  27174171E2C3F03D CRC64;

  Query Match             13.6%;  Score 322.5;  DB 10;  Length 623;
  Best Local Similarity   43.4%;  Pred. No. 2.9e-18;
  Matches   76;  Conservative   28;  Mismatches   38;  Indels   33;  Gaps    6;

Qy        237 LTSEAVVPATKTRKLSAERSEARSHLL-LQKRQFYHSHRVQPMALEQVMSDRDSEDEVD- 294
              ||:|| ||        |:||:| || | | |||||||   ||::||||||||||::||
Db          6 LTTEAKVP--------AKRSKATSHYLPLHKRQFYHSRTGQPLSLEQVMSDRDSENDVDK 57

Qy        295 -DDVADFEDRQMLDDFVDVNK-DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKE 352
               || |  |: |||:  :| |:   ::|: ||||||::||::|| || |||||||| : :|
Db         58 NDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAHIPWACEAFSRLHLQE 117

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              |    ||      | |    ::| :              ::: |:|    : :|
Db        118 LRSNLSL---------DLGNTETTETAPV-----------ATEATTHGEQDGID 151


RESULT 11
O23526
ID   O23526      PRELIMINARY;      PRT;   295 AA.
AC   O23526;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
```

```
DE   Hypothetical 33.6 kDa protein.
GN   AT4G16850.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA   Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA   Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA   Jones J., Palme K., Ansorge W., Delseny M., Bancroft I., Mewes H.W.,
RA   Schueller C., Chalwatzis N.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97342; CAB10459.1; -.
DR   EMBL; AL161545; CAB80956.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   295 AA;  33564 MW;  F22DBD5E82031423 CRC64;

  Query Match             11.2%;  Score 266;  DB 10;  Length 295;
  Best Local Similarity   53.4%;  Pred. No. 5.3e-14;
  Matches   70;  Conservative    3;  Mismatches   30;  Indels   28;  Gaps    4;

Qy        278 MALEQVMSDRDSEDEVDDDVADFEDRQ------------------MLDDFVDVNKDEKQ 318
              |||||||||||||||||||||||||||                  ||||||||||||||
Db          1 MALEQVMSDRDSEDEVDDDVADFEDRQVFHDFFLRSLSRQQKMVYDMLDDFVDVNKDEKQ 60

Qy        319 FMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
              |||||||||||||   : ||     ||       ||   ||     |   :| ||
Db         61 FMHLWNSFVRKQR---NKHILLPIFAFIAIPLAALH--LSL----TLTSFRLKNHVFRLE 111

Qy        379 ATINNCNTILE 389
              |  |  :|  |
Db        112 ALANVVHTRFE 122


RESULT 12
Q96BD9
ID   Q96BD9      PRELIMINARY;      PRT;   739 AA.
AC   Q96BD9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Joined to JAZF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015704; AAH15704.1; -.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   739 AA;  83054 MW;  A8830EBC3FD38D56 CRC64;

  Query Match             9.3%;  Score 221;  DB 4;  Length 739;
  Best Local Similarity   21.1%;  Pred. No. 9e-10;
  Matches   84;  Conservative   64;  Mismatches  151;  Indels  100;  Gaps   12;

Qy         58 RKRK-----SRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
              || |     :|    : : :   ||| |:|| |:|  ||:|:: |     |  ||   |
Db        413 RKEKDTPNENRQKLRIFYQFLYNNNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHS 472

Qy        113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN 168
              | | :    :   ::||:         |  |       |    |:  || |    ||
Db        473 RFIFNYVYHPKGARIDVSI--------NECYDG------SYAGNPQDIHRQPGFAFSRNG 518

Qy        169 -TRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHS 227
               :|  :  : :  |  |  : :      ||
Db        519 PVKRTPITHILVCRPKRTKASMSEFLESEDG---------------------------- 549

Qy        228 SLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDR 287
                                  ::  :|: :  |     | ::||    |:  :::  :
Db        550 ----------------------EVEQQRTYSSGH----NRLYFHSDTCLPLRPQEM--EV 581

Qy        288 DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSR 347
              |||||  |  :    :   :::| |||: ||: | |||  | |   |||  :: ||  |
Db        582 DSEDEKDPEWLREKTITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVE 641

Qy        348 FYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              | ::: :  :|   : | |: : :  |:   :|:   | |  :             ::
Db        642 NYGQKIIK-KNLCRNFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLE 689

Qy        408 RPSDSNTNNNNIVDHPN-------DINNKNNVDNKDNNS 439
              :   ::  |  | :  |       :||:|      |: |
Db        690 KGESASPANEEITEEQNGTANGFSEINSKEKALETDSVS 728


RESULT 13
Q15022
ID   Q15022      PRELIMINARY;      PRT;   803 AA.
AC   Q15022;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   KIAA0160 protein (Fragment).
GN   KIAA0160.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127530; PubMed=8590280;
RA   Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. IV.
RT   The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 2:167-174(1995).
DR   EMBL; D63881; BAA09931.1; -.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   803 AA;  89963 MW;  CDFB901A35F29A7C CRC64;

  Query Match             9.3%;  Score 221;  DB 4;  Length 803;
  Best Local Similarity   21.1%;  Pred. No. 1e-09;
  Matches   84;  Conservative   64;  Mismatches  151;  Indels  100;  Gaps   12;

Qy         58 RKRK-----SRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
              || |     :|    : : :   ||| |:|| |:|  ||:|:: |     |  ||   |
Db        477 RKEKDTPNENRQKLRIFYQFLYNNNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHS 536

Qy        113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN 168
              | | :    :   ::||:         |  |       |    |:  || |    ||
Db        537 RFIFNYVYHPKGARIDVSI--------NECYDG------SYAGNPQDIHRQPGFAFSRNG 582

Qy        169 -TRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHS 227
               :|  :  : :  |  |  : :      ||
Db        583 PVKRTPITHILVCRPKRTKASMSEFLESEDG---------------------------- 613

Qy        228 SLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDR 287
                                  ::  :|: :  |     | ::||    |:  :::  :
Db        614 ----------------------EVEQQRTYSSGH----NRLYFHSDTCLPLRPQEM--EV 645
```

```
Qy        288 DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSR 347
              ||||| | :    :    :::| |||: ||: | |||  | |   |||  :: ||  |
Db        646 DSEDEKDPEWLREKTITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVE 705

Qy        348 FYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
               | ::: :  :|   : | |: : :  |:   :|:   | |   :             ::
Db        706 NYGQKIIK-KNLCRNFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLE 753

Qy        408 RPSDSNTNNNNIVDHPN-------DINNKNNVDNKDNNS 439
              :   ::  |  | :  |       :||:|     |: |
Db        754 KGESASPANEEITEEQNGTANGFSEINSKEKALETDSVS 792


RESULT 14
Q8T9D8
ID   Q8T9D8      PRELIMINARY;      PRT;   855 AA.
AC   Q8T9D8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   SD04959p.
GN   SU(Z)12.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069809; AAL39954.1; -.
SQ   SEQUENCE   855 AA;  95316 MW;  C10FFC6013954105 CRC64;

  Query Match             9.2%;  Score 219;  DB 5;  Length 855;
  Best Local Similarity   22.6%;  Pred. No. 1.6e-09;
  Matches   74;  Conservative   47;  Mismatches  113;  Indels   94;  Gaps    8;

Qy         68 VVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTV 127
              :|:|:   ||| |:|| :: :||:| : |     |  ||   |  | | ::       :
Db        391 IVYNFMYSNNTRQQTEYTQELNCPWCGLDCLRLYALLKHLKLCHARFNFTYQPAGSGARI 450

Qy        128 NVSVKLNSFIFEEEGSDDDKFEP----FSLCSKPRKRRQRGGRN--NTRRLKVC---FLP 178
              :|::  :::     ||  |   |    |:    | :|          || | |   ||
Db        451 DVTIN-DAYDGSYAGSPYDLAGPSGSSFARTCGPVRRTSVTSLMVCRPRRQKTCLDEFLE 509

Qy        179 LDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILT 238
              ||   ::|
Db        510 LDEDEISN---------------------------------------------------- 517

Qy        239 SEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVA 298
                                ||::    | ::|:    |:  :::  | ||| | |
Db        518 ------------------QRSYITGHNRLYHHTETCLPVHPKEL--DIDSEGESDPLWL 556

Qy        299 DFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAF-----SRFYEKEL 353
               :  ||:|:| |||: ||: | |||  | :   : |  :  ||| |     :    | |
Db        557 RQKTIQMIDEFSDVNEGEKELMKLWNLHVMRHGFVGDCQLPIACEMFLDAKGTEIVRKNL 616

Qy        354 HRYSSLFWCWRLFLIKLWNHGLVDSATI 381
              :|   |  |       |:::||: : |:
Db        617 YRNFILHMC------SLFDYGLIAAETV 638


RESULT 15
Q9NJG9
ID   Q9NJG9      PRELIMINARY;      PRT;   900 AA.
AC   Q9NJG9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Su(z)12.
GN   SU(Z)12 OR CG8013.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Birve A., Rasmuson-Lestander A., Larsson J.;
RT   "Suppressor of zeste 12.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF149047; AAF73149.1; -.
DR   FlyBase; FBgn0020887; Su(z)12.
DR   InterPro; IPR000822; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ   SEQUENCE   900 AA;  100103 MW;  53BA0D83C49EC92F CRC64;

  Query Match             9.2%;  Score 219;  DB 5;  Length 900;
  Best Local Similarity   22.6%;  Pred. No. 1.7e-09;
  Matches   74;  Conservative   47;  Mismatches  113;  Indels   94;  Gaps    8;

Qy         68 VVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTV 127
              :|:|:   ||| |:|| :: :||:| : |     |  ||   |  | | ::       :
Db        391 IVYNFMYSNNTRQQTEYTQELNCPWCGLDCLRLYALLKHLKLCHARFNFTYQPAGSGARI 450

Qy        128 NVSVKLNSFIFEEEGSDDDKFEP----FSLCSKPRKRRQRGGRN--NTRRLKVC---FLP 178
              :|::  :::     ||  |   |    |:    | :|          || | |   ||
Db        451 DVTIN-DAYDGSYAGSPYDLAGPSGSSFARTCGPVRRTSVTSLMVCRPRRQKTCLDEFLE 509

Qy        179 LDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILT 238
              ||   ::|
Db        510 LDEDEISN---------------------------------------------------- 517

Qy        239 SEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVA 298
                                ||::    | ::|:    |:  :::  | ||| | |
Db        518 ------------------QRSYITGHNRLYHHTETCLPVHPKEL--DIDSEGESDPLWL 556

Qy        299 DFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAF-----SRFYEKEL 353
               :  ||:|:| |||: ||: | |||  | :   : |  :  ||| |     :    | |
Db        557 RQKTIQMIDEFSDVNEGEKELMKLWNLHVMRHGFVGDCQLPIACEMFLDAKGTEIVRKNL 616

Qy        354 HRYSSLFWCWRLFLIKLWNHGLVDSATI 381
              :|   |  |       |:::||: : |:
Db        617 YRNFILHMC------SLFDYGLIAAETV 638


Search completed: June 13, 2003, 15:45:22
Job time : 85 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:17:53 ; Search time 1928 Seconds
(without alignments)
3738.066 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144851_16490/app_query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_1906_@runat_13062003_144851_16490 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		778	32.7	672	10	AV822548	AV822548 AV822548
2		760.5	32.0	1128	11	AY104964	AY104964 Zea mays
3		568	23.9	680	10	AW038171	AW038171 EST279828
4	c	553	23.3	769	14	BQ505017	BQ505017 EST612432
5		545	22.9	846	12	BG648271	BG648271 EST509890
6		528	22.2	480	9	AJ469021	AJ469021 AJ469021
7		510.5	21.5	561	14	BQ740672	BQ740672 saq50e05.
8		509	21.4	600	13	BI479743	BI479743 WHE3451_E
9	c	506	21.3	391	10	AV817525	AV817525 AV817525
10		498.5	21.0	569	14	BQ611550	BQ611550 sap61f09.
11		456.5	19.2	474	10	AW234600	AW234600 sf17b08.y
12		450	18.9	467	12	BG580495	BG580495 EST482220
13		448	18.8	405	12	BG314044	BG314044 WHE2467_D
14	c	444.5	18.7	662	14	BQ279630	BQ279630 1091030A1
15		442.5	18.6	612	10	BE203328	BE203328 EST403350
16		437	18.4	444	12	BG551103	BG551103 sad33c03.
17		436	18.3	554	13	BI784785	BI784785 saf93c12.
18		422	17.7	429	14	BQ459233	BQ459233 HA01O05r
19		420	17.7	546	13	BI321750	BI321750 saf98e10.
20		396.5	16.7	576	10	AW076157	AW076157 614063H03
21		394	16.6	364	9	AI164598	AI164598 A065P57U
22		386	16.2	732	14	BQ841916	BQ841916 WHE2986_G
23	c	375	15.8	561	14	BQ162202	BQ162202 WHE0751_H
24		374	15.7	644	13	BJ453006	BJ453006 BJ453006
25	c	372	15.6	757	13	BJ460553	BJ460553 BJ460553
26		367.5	15.5	445	10	AW202144	AW202144 sf12d08.y
27		362.5	15.2	594	10	AW000259	AW000259 614009G07
28	c	347	14.6	519	14	BQ168835	BQ168835 WHE2467_D
29		343.5	14.4	326	9	AI163743	AI163743 A047P36U
30		329	13.8	675	13	BI932726	BI932726 EST552615
31		306	12.9	727	14	BQ801572	BQ801572 WHE2816_A
32		304.5	12.8	560	9	AI795561	AI795561 614009G07
33	c	304	12.8	526	17	CNS00WYB	AL094193 Arabidops
34	c	303	12.7	586	14	BQ243123	BQ243123 TaE15019A
35		302.5	12.7	529	12	BG459005	BG459005 947052F03
36		299.5	12.6	515	12	BF481866	BF481866 FM1_22_D0
37	c	274	11.5	677	14	BQ659737	BQ659737 HF01I23w
38	c	271	11.4	322	10	AW309408	AW309408 sf17b08.x
39	c	269	11.3	236	14	BQ904049	BQ904049 Ta03_18d1
40	c	268	11.3	804	17	BH587461	BH587461 BOHJH08TF
41	c	267	11.2	553	12	BG909756	BG909756 TaLr1108b
42	c	267	11.2	553	14	BQ620168	BQ620168 TaLr1108B
43	c	257	10.8	523	13	BI992698	BI992698 1020066F1
44	c	256.5	10.8	632	12	BE821501	BE821501 GM700014B
45		254.5	10.7	416	10	AW432863	AW432863 sh99a03.y

ALIGNMENTS

RESULT 1
AV822548
LOCUS AV822548 672 bp mRNA linear EST 01-APR-2002
DEFINITION AV822548 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-08-K19 5', mRNA sequence.
ACCESSION AV822548
VERSION AV822548.1 GI:19864597
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 672)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,

```
            Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
            and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998).This clone is in a
            modified pBluescript vector as a SstI/XhoI insert. Please visit our
            web site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .672
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL05-08-K19"
                     /clone_lib="RAFL5"
                     /dev_stage="rosette plants"
                     /lab_host="SOLR"
                     /note="Site_1: SstI; Site_2: XhoI; subjected to
                     dehydration-treated(1,2,5,10,24 hr)"
BASE COUNT      197 a    116 c    143 g    216 t
ORIGIN

Alignment Scores:
Pred. No.:                6.11e-80      Length:          672
Score:                    778.00        Matches:         144
Percent Similarity:       99.32%        Conservative:    1
Best Local Similarity:    98.63%        Mismatches:      1
Query Match:              32.72%        Indels:          0
DB:                       10            Gaps:            0

US-09-890-220-2 (1-445) x AV822548 (1-672)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAsp 146
              |||||||||||||||||
Db        653 GAAGGAAGTGATGATGAT 670

RESULT 2
```

```
AY104964
LOCUS       AY104964                1128 bp    mRNA    linear   HTC 25-MAY-2002
DEFINITION  Zea mays PCO105783 mRNA sequence.
ACCESSION   AY104964
VERSION     AY104964.1  GI:21208042
KEYWORDS    HTC.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 1128)
  AUTHORS   Coe,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .1128
                     /organism="Zea mays"
                     /db_xref="MaizeDB:636128"
                     /db_xref="taxon:4577"
                     /clone="PCO105783"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      355 a    199 c    242 g    332 t
ORIGIN

Alignment Scores:
Pred. No.:                1.48e-77      Length:          1128
Score:                    760.50        Matches:         157
Percent Similarity:       64.35%        Conservative:    47
Best Local Similarity:    49.53%        Mismatches:      80
Query Match:              31.98%        Indels:          33
DB:                       11            Gaps:            7

US-09-890-220-2 (1-445) x AY104964 (1-1128)

Qy         86 GluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeuGlnPhe 105
              ||||||   ||||||||||||||||   :::   ||||||||||||||||||||
Db         12 GAAGATTTCTCTTGTCCATTTTGCTATGTACGATGTGGAAGCTTCAAGGGTCTAGGATGC 71

Qy        106 HisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGluTyrGln 125
              ||||||||||||||||||||||||||   :::|||||||   :::   ||||||||||||
Db         72 CATTTAAACTCATCACATGATCTATTCCACTATGAGTTTTGGATATCTGAAGAGTACCAG 131

Qy        126 ThrValAsnValSerValLysLeuAsnSerPhe---------IlePheGluGluGluGly 142
                 ||||||||||||:::|||   :::::::::          :::|||   |||
Db        132 GTTGTTAATGTTAGTCTGAAGGCTGATGCTTGGAGAACAGAGCTTTTTGCGGAG------ 185

Qy        143 SerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgArgGln 162
                 |||   :::   :::   ||||||      |||:::   :::||||||||:::
Db        186 GGCGTTGATCCAAGGCATCAAACATTTTCTTATCGCTCAAGGTTTAAGAAGCGTAGACGA 245

Qy        163 ---ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSer 181
                 :::                  |||   :::                :::         |||
Db        246 TCAAAGAACACAATGGAGAAAATCAGGCATGTACACTCACATATTATGGAATCAGGTTCA 305

Qy        182 ProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGly 201
              |||           |||:::|||:::                        |||      |||
Db        306 CCTGAAGATGAGGCAGGATCTGAGGACAACTTTGTG---------CAAGGGGAGAATGGG 356
```

```
Qy        202 LeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIlePro 221

Db        356 ------------------------------------------------------------ 356

Qy        222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal---------IleLeuThr 238
                  ::::::|||::::::|||:::|||             :::          |||::: 
Db        357 ACTTCTGTAGCAAATGCTTCGATTGATCCTGCTCAATCTTTACATGGCAGCAATCTTTCA 416

Qy        239 SerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAla 258
                     |||:::         ||||||||||||||||||   ||||||||||:::    
Db        417 CCACCAACAGTACTACAGTTTGGGAAGACAAGGAAGCTATCT---GAGAGATCTGACCCT 473

Qy        259 ArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMet 278
                  |||:::      ||||||||||||||||||:::||||||||||||   |||||||||
Db        474 AGAAATCGGCAACTCCTGCAAAAACGACAGTTCTTCCATTCTCACAGGGCGCAGCCAATG 533

Qy        279 AlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAla 298
                  ||||||||||||   ||||||||||||||||||||||||||||||||||:::|||  
Db        534 CAACTGGAGCAAGTGTTCTCGGACCGTGATAGTGAAGATGAAGTTGATGATGATATTGCT 593

Qy        299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGln 318
                  ||||||||||||||:::||||||||||||||||||||||   ||||||||||||   
Db        594 GACTTCGAGGATAGAAGAATGCTTGATGATTTTGTTGATGTTACGAAAGATGAAAAACTT 653

Qy        319 PheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIle 338
                  |||||||:::|||||||||||||||||||||||||||||:::|||||||||||||||
Db        654 ATTATGCATATGTGGAATTCATTTGTTCGAAAACAAAGAGTGTTAGCTGATGGTCATATA 713

Qy        339 SerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSer 358
                  ||||||||||||||||||||:::   :::   ::::::|||   :::      ::: 
Db        714 CCTTGGGCCTGCGAGGCATTCTCCCAGTTGCATGGACGACAACTTATACAAAATCCTGCT 773

Qy        359 LeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSer 378
                  |||   |||   ||||||   |||:::||||||||||||||||||   ::::::|||::: 
Db        774 CTGCTGTGGGGTTGGCGTTTCTTCATGATTAAACTTTGGAACCATAACATTTTAGATGCC 833

Qy        379 AlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
                  |||:::|||   ||||||||||:::|||:::      :::      |||          
Db        834 CGCACTATGAACACATGCAATACAGTCCTTCAAATTTTACAAGAAGAAAGC 884

RESULT 3
AW038171
LOCUS       AW038171                680 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST279828 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
            clone cLET1N5, mRNA sequence.
ACCESSION   AW038171
VERSION     AW038171.1  GI:5896925
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
            Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
            Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni
            ,J.
  TITLE     Generation of ESTs from tomato leaf tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Lycopersicon esculentum"
                     /cultivar="Rio Grande PtoR"
                     /db_xref="taxon:4081"
                     /clone="cLET1N5"
                     /clone_lib="tomato mixed elicitor, BTI"
                     /tissue_type="leaf"
                     /dev_stage="4-6 week old plants"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLET - Inoculated with a variety of disease response
                     elicitors. Plants  exposed to 2,6 dichloroisonicotinic
                     acid, BTH, jasmonic acid, ethylene,  fenthion, EIX,
                     okadaic acid, or systemin prior to tissue harvest.  EcoRI
                     site was destroyed during cloning."
BASE COUNT      192 a    120 c    164 g    204 t
ORIGIN

Alignment Scores:
Pred. No.:              1.95e-55        Length:         680
Score:                  568.00          Matches:        107
Percent Similarity:     84.35%          Conservative:   17
Best Local Similarity:  72.79%          Mismatches:     21
Query Match:            23.89%          Indels:         2
DB:                     10              Gaps:           2

US-09-890-220-2 (1-445) x AW038171 (1-680)

Qy        254 GluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
                  ||||||||||:::   |||:::      ||||||||||||||||||||:::|||||||||
Db          4 GAGCGCTCTGATCCCAGAAATCGTGCACTCCTGCAAAAAAGGCAATTCTTTCATTCTCAT 63

Qy        274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
                  |||   ||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         64 AGGGCCCAGCCCATGGCACTGGAGCAAGTTTTATCGGACCGAGACAGTGAGGATGAAGTT 123

Qy        294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
                  ||||||||||||||||   |||||||||:::|||||||||||||||||||||||||
Db        124 GATGATGATGTTGCAGATCTTGAAGATCGAAGGATGCTTGATGATTTTGTGGATGTGACC 183

Qy        314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
                  ||||||||||||||.|   ||||||||||||||||||||||||||||||||||||:::
Db        184 AAAGATGAAAAGCAAGTGATGCATCTGTGGAACTCATTTGTTAGAAAGCAAAGGGTGTTG 243

Qy        334 AlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyr---GluLysGlu 352
                  ||||||||||||||   ||||||||||||||||||||:::   :::   ::::::
Db        244 GCAGATGGTCATATCCCTTGGGCATGTGAGGCCTTTTCAAAGCTGCATGGTCAGAGGTTT 303

Qy        353 LeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
                        :::::::|||      ||||||||||||||:::::|||||||||||
Db        304 GCCCAAGCACCAGCCACACTATGCAGGTGTTGGAGATTATTCATGATGAAGCTGTGGAAC 363

Qy        373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
                  ||||||||||||||:::   ||||||||||||||||||   |||||||||      :::
Db        364 CATGGCCTTGTTGATGCGCGTACAATTAACAATTGTAACCTAATATTAGAGCAGTTCCAA 423

Qy        393 AsnSer---SerAspThrThr 398
                  |||      ||||||:::|||
Db        424 AACCAATACAGTGATTCTACT 444

RESULT 4
BQ505017/c
LOCUS       BQ505017                769 bp    mRNA    linear   EST 22-JUL-2002
DEFINITION  EST612432 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
            3' end, mRNA sequence.
ACCESSION   BQ505017
VERSION     BQ505017.2  GI:21920982
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
```

```
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21363886.
            Other_ESTs: EST612431
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMGB40"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating eyes
                     , tubers, or roots."
BASE COUNT      239 a    176 c    140 g    214 t
ORIGIN

Alignment Scores:
Pred. No.:                1.32e-53        Length:          769
Score:                    553.00          Matches:         113
Percent Similarity:       79.39%          Conservative:    18
Best Local Similarity:    68.48%          Mismatches:      31
Query Match:              23.25%          Indels:          4
DB:                       14              Gaps:            2

US-09-890-220-2 (1-445) x BQ505017 (1-769)

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
              |||          :::                 |||:::|||||||||||||   |||||||||
Db        769 CTTGCACCCCCCAGCACTGCTTCAGTTGCCAAGTCAAGAAAGTTATCAGTTGAGCGCTCT 710

Qy        257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
                  |||:::      |||||||||||||||||   :::|||||||||||   |||
Db        709 ATCCCCAGAAATCGTGCACTCCTGCAAAAAAGGCAATC-TTTCATTCTCATAGGGCCCAG 651

Qy        277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
              |||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||
Db        650 CCCATGGCACTGGAGCAAGTTTTATCAGATCGAGACAGTGAGGATGAAGTTGATGATGAT 591

Qy        297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
              |||||||||   |||||||||:::|||||||||||||||||||||||   |||||||||
Db        590 GTTGCAGATCTTGAAGATCGAAGGATGCTTGATGATTTTGTGGATGTGACCAAAGATGAA 531

Qy        317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
              ||||||   ||||||||||||||||||||||||||||||||||||:::|||||||||
Db        530 AAGCAAATGATGCATCTGTGGAACTCATTTGTTAGAAAGCAAAGGGTGTTGGCAGATGGT 471

Qy        337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu---HisArg 355
              ||||||   |||||||||||||||||||:::   ::::::   ::::::
Db        470 CACATCCCTTGGGCATGTGAGGCCTTTTCAAAGCTGCATGATTCAGATGTTTGCCCAAGC 411

Qy        356 TyrSer-SerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLe 375
              ||| ||||||      |||||||||||||||::::::|||||||||||||||||||||
Db        410 ACCAGCCTCGTTGTGCAGGTGTTGGAGATTATTCATGATGAAGTTGTGGAACCATGGCCT 351

Qy        375 uValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn---Se 394
              |||||||:::   ||||||||||||||||||   |||||||||      ::::::
Db        350 TGTTGATGCGCGTACAATTAACAATTGTAACCTAATATTAGAGCAGTTCCAAAGCCAAGA 291

Qy        394 rSerAspThrThr 398
```

```
                  ||||||:::|||
Db        290 CAGTGATTCTACT 278

RESULT 5
BG648271
LOCUS       BG648271                846 bp    mRNA    linear   EST 24-APR-2001
DEFINITION  EST509890 HOGA Medicago truncatula cDNA clone pHOGA-19E20 5' end,
            mRNA sequence.
ACCESSION   BG648271
VERSION     BG648271.1  GI:13783383
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
            Utterback,T., Cho,J. and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula treated with
            oligogalacturonides of DP 6-20
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Michael G. Hahn
            Complex Carbohydrate Research Center
            University of Georgia
            220 Riverbend Road, Athens, GA 30602-4712, USA
            Tel: 706-542-4457
            Fax: 706-542-4412
            Email: hahn@ccrc.uga.edu
            G391877e TIGR sequence name: MTMCJ34TK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Medicago truncatula"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pHOGA-19E20"
                     /clone_lib="HOGA"
                     /tissue_type="3 day old seedling roots"
                     /dev_stage="24 hours after treatment in the dark at 26 C
                     with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
                     presence of 100 ug/ml Gentamicin"
                     /lab_host="XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in SOLR cells."
BASE COUNT      248 a    161 c    185 g    252 t
ORIGIN

Alignment Scores:
Pred. No.:                1.31e-52        Length:          846
Score:                    545.00          Matches:         123
Percent Similarity:       58.66%          Conservative:    26
Best Local Similarity:    48.43%          Mismatches:      68
Query Match:              22.92%          Indels:          37
DB:                       12              Gaps:            6

US-09-890-220-2 (1-445) x BG648271 (1-846)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||||||||            :::|||   ||||||   |||:::   ||||||
Db         52 ATGTGCCGGCAAAATTCACCGGTGCATAATTCCGGTGAAGAAGAAATTACAGCTGATGAA 111

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||      |||||||||   :::   |||||||||
Db        112 AGTCTTTTGATTTATTGCAAACCTGTTGAATTCTACAATATTCTCTATCGTCGTTCTCTT 171
```

```
Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
Db        172 CACAATCCTACTTTTCTTAAGAGATGTTTGCGCTATAAAATAAAAGCAAAGCAAAAAAGG 231

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
Db        232 AGGTTGCGAGCA---GGAATTGTGGTTTTCAACTATAGGGACTGTTACAATGCGCTTCGA 288

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
Db        289 AAGACTGAAGTGACTGAAGACTTTTCTTGTCCATTTTGCTTGATGCAGTGTGCGAGCTTT 348

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
Db        349 AAGGGTTTGCGATTTCATCTTTGTGCATCACATGATCTATTCAACTTTGAGTTCTGGGTC 408

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
Db        409 ACCAATGATTACCAAGCAGTGAACGTCTCTGTAAAAACTGAC------ATATTGAGATCA 462

Qy        141 GluGlySerAspAsp---------AspLysPheGluProPheSerLeuCysSerLysPro 157
Db        463 GAGAATGTTGCTGATGGAGTAATTCCACAATCACAAACCTTCTTCTTCTGTTCAAGATCT 522

Qy        158 ArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
Db        523 CGAAAGCGTAGACCAAGAGGCTCCGTTCAAAATGGGAAGCATGCCAATGTAAATTTCCTG 582

Qy        178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIle--------------- 192
Db        583 AAGTTGGATTCACCA---------GAAGGCATACAGAATGGGGTTCTACAGAAAGACAAC 633

Qy        193 ----------------------------------------------------------ThrLeu 194
Db        634 GATATCCTGTCCAGCAAAGGGGAGAATATGTCAAGAGCATCTCGTAGTGAGAAGAATTTG 693

Qy        195 LeuAsnAspGlyAsnArg------------GlyLeuGlyTyrProGluAlaThrGluLeu 210
Db        694 CAGAATGAGGGAAATGGTTGGGGTAAATTTGGTCCTGATCATCCTAGTGCCATGGATTAC 753

Qy        211 AlaGly-GlnPheGluMetThrSerAsnIleProProAla 223
Db        754 GCGGGACATGTTGGAATCTAGTTTTAACATTTCCAGTGCT 793
```

RESULT 6
AJ469021
LOCUS AJ469021 480 bp mRNA linear EST 24-MAY-2002
DEFINITION AJ469021 S00008 Hordeum vulgare cDNA clone S0000800053C04F1, mRNA sequence.
ACCESSION AJ469021
VERSION AJ469021.1 GI:21184977
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 480)
AUTHORS Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
TITLE Barley EST's
JOURNAL Unpublished (2002)
COMMENT Contact: Schulman AH
Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014, Finland.
FEATURES Location/Qualifiers
source 1. .480
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="S0000800053C04F1"
/clone_lib="S00008"
/tissue_type="Callus"
/note="Callus K19"
BASE COUNT 138 a 95 c 111 g 136 t
ORIGIN

Alignment Scores:

Pred. No.:	5.31e-51	Length:	480
Score:	528.00	Matches:	95
Percent Similarity:	79.14%	Conservative:	15
Best Local Similarity:	68.35%	Mismatches:	29
Query Match:	22.20%	Indels:	0
DB:	9	Gaps:	0

US-09-890-220-2 (1-445) x AJ469021 (1-480)

```
Qy        257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
Db          1 GATCCCAGAAACCGACAACTCCTGCAGAAACGTCAGTTTTTCCATTCTCACAGGGCACAG 60

Qy        277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
Db         61 CCAATGGCACTGGAACAAGTTTTCTCGGATCGTGATAGTGAAGATGAAGTTGATGATGAC 120

Qy        297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
Db        121 ATCGCAGATTTTGAAGATAGACGGATGCTTGATGATTTTGTTGATGTCACGAATGATGAG 180

Qy        317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
Db        181 AAACTTATTATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTGCTAGCTGATGGT 240

Qy        337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
Db        241 CATATTCCTTGGGCCTGTGAGGCATTCTCCCGGCTTCATGGAAAACATCTTGTACAGAAT 300

Qy        357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
Db        301 CCTCCTCTACTATGGAGCTGGCGTTTCCTTATGATTAAACTGTGGAACCACAGTCTATTA 360

Qy        377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
Db        361 GATGCCCGTGCCATGAATGTCTGCGGCACAATTCTTCAAGGCTACCAAAATGAAAGC 417
```

RESULT 7
BQ740672
LOCUS BQ740672 561 bp mRNA linear EST 17-JUL-2002
DEFINITION saq50e05.y1 Gm-c1076 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1076-3274 5' similar to TR:Q9ZNT9 Q9ZNT9 FERTILIZATION-INDEPENDENT SEED 2 PROTEIN. ;, mRNA sequence.
ACCESSION BQ740672
VERSION BQ740672.1 GI:21887459
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 561)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

```
          This clone is available through: ResGen, Invitrogen Corp. 2130
          South Memorial Parkway Huntsville, AL 35801 For further information
          call: (800)-533-4363 or contact: ccu@resgen.com web site:
          www.resgen.com
          Seq primer: -40RP from Gibco
          High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .561
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1076-3274"
                     /clone_lib="Gm-c1076"
                     /tissue_type="wounded cotyledons"
                     /dev_stage="11 day old seedlings"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from mRNA isolated
                     from 11 day old seedlings treated with that were treated
                     with 2 ugs/ml of a crude glucan elicitor preparation
                     isolated from the mycelial walls of Phytophthora sojae.
                     The  library was prepared using the Stratagene pBluescript
                     II SK(+) library  construction kit. Complementary DNA was
                     synthesized from mRNA using a primer  consisting of a
                     poly(dT) sequence with an XhoI restriction site. EcoRI
                     adaptors were ligated to the blunt-ended cDNA fragments
                     followed by XhoI  digestion. The cDNA fragments were
                     directionally cloned into the EcoRI-XhoI  restriction site
                     of the pBluescript vector. The ligated cDNA fragments were
                     transformed into E.coli ElectroMax DH10B host cells. Plant
                     material was provided by Michael G. Hahn (Complex
                     Carbohydrate Research Center, University of Georgia) and
                     the library was constructed by Anu Khanna (Lila Vodkin lab
                     , University of Illinois)."
BASE COUNT      166 a     91 c    123 g    181 t
ORIGIN

Alignment Scores:
Pred. No.:              7.41e-49        Length:         561
Score:                  510.50          Matches:        93
Percent Similarity:     77.78%          Conservative:   19
Best Local Similarity:  64.58%          Mismatches:     27
Query Match:            21.47%          Indels:         5
DB:                     14              Gaps:           1

US-09-890-220-2 (1-445) x BQ740672 (1-561)

Qy        280 LeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAsp 299
              ||||||||||||::||||||||||||||||||||||||||||||||:::|||||||
Db          2 CTAGAACAAGTGTTATCAGACCGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGAT 61

Qy        300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPhe 319
              |||||||||:::|||||||||||||||||||||||:::|||||||||||||||
Db         62 CTTGAAGATAGAAGGATGCTTGACGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTC 121

Qy        320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
              ||||||||||||||||||||:::|||||||||||||||:::|||||||||||||:::
Db        122 ATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGGGTGCTAGCAGATGGTCATGTTCCG 181

Qy        340 TrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeu 359
              |||||||||||||||||||:::   :::   |||||||||             :::|||
Db        182 TGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAAGAGCTGATCTCATCTCCAGCTTTA 241

Qy        360 PheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAla 379
              ||||||||||||||||||||::||||||||||||||||||||||||:::|||:::
Db        242 TTTTGGTGTTGGAGGTTATTCATGATCAAACTTTGGAATCATGGTCTTCTTGATGCCTGT 301

Qy        380 ThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 399
              |||:::||||||||:::   :::|||::::::  ||||||          |||
Db        302 ACAATGAACAACTGTAGCATAGTATTAGATAGTTACAGAAATGAGGGATCGGGTACAAGA 361

Qy        400 ThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIle 419
              |||                 ||||||     :::     ||||||  ||||||
Db        362 AAAAAT--------------TGAAGACCATAAGGAAACTGACTCAACAACAAGAATATA 406

Qy        420 ValAspHisPro 423
                       |||
Db        407 CAGTCTTCTCCG 418

RESULT 8
BI479743
LOCUS       BI479743                 600 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  WHE3451_E10_I19ZS Wheat pre-anthesis spike cDNA library Triticum
            aestivum cDNA clone WHE3451_E10_I19, mRNA sequence.
ACCESSION   BI479743
VERSION     BI479743.1  GI:15323794
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE3451_E10_I19"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /tissue_type="Spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse.  Whole spike with awns trimmed,  white, green
                     and yellow anther were collected and total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the cDNA  clones were in vivo excised to give pBluescript
                     phagemids in the TJ  Close lab (Choi, Close, Fenton) at
                     the University of California,  Riverside.  Plasmid DNA
                     preparations and DNA sequencing were performed  in the OD
                     Anderson lab (all other authors)."
BASE COUNT      165 a    126 c    143 g    166 t
ORIGIN

Alignment Scores:
Pred. No.:              1.23e-48        Length:         600
Score:                  509.00          Matches:        102
Percent Similarity:     62.18%          Conservative:   18
Best Local Similarity:  52.85%          Mismatches:     30
Query Match:            21.40%          Indels:         44
DB:                     13              Gaps:           2

US-09-890-220-2 (1-445) x BI479743 (1-600)

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
              ||||||         |||:::         ||||||||||||||||||||||||||||||
Db         21 CTTACACCACCGACAGTACTAGAGTTTGGGAAGACAAGGAAACTATCTGCGGAGCGAAGT 80

Qy        257 GluAlaArgSer---------------------------------------------- 260
```

```
                :::   |||
Db          81 GATCCCAGAAA-GTAGCCCTTCTGTTTACATACTTCGCTGATATCATGCTAAGCATTGTA 139

Qy         261 ------------------------HisLeu------------------------------ 262
                                       ||||||
Db         140 TATATAGCTGTTTGTTTGTCCAGACATTTACAATAGAAAAATGTTAAACCAACTGTTTTT 199

Qy         263 --------------------------LeuLeuGlnLysArgGlnPheTyrHisSerHis 273
                                         ||||||||||||||||||||:::|||||||||
Db         200 ATGTGTATATTATGTGGCAGCCGACAACTTCTGCAGAAACGTCAGTTCTTCCATTCTCAC 259

Qy         274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
               |||   |||||||||||:::||||||||||:::|||||   ||||||||||||||||||
Db         260 AGGGCACAGCCAATGGCAGTGGAACAAGTTTTGTCAGATCATGACAGCGAGGATGAAGTT 319

Qy         294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
               |||||||||:::||||||   ||||||||||:::||||||   ||||||:::||||||
Db         320 GACGATGACATTGCCGACTTGGAAGATAGACGGATGCTTGCTGATTTTCTTGATGTCACG 379

Qy         314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
               ||||||||||||      ||||||:::||||||||||||:::||||||||||||||:::
Db         380 AAAGATGAGAAGCTTATTATGCATATGTGGAATTCATTTATTCGGAAACAAAGGGTGCTA 439

Qy         334 AlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353
               ||||||||||||||   ||||||||||||   ||||||||   :::      :::|||
Db         440 GCTGATGGGCATATACCTTGGGCCTGTGAGGGCTTCTCCCGGCTTCATGGACCGCAGCTT 499

Qy         354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373
                  :::         |||   ||||||||||||||   :::|||||||||||||||||
Db         500 GTACAAAACCCTCCTCTGCTCTGGTGCTGGCGTTTGGTGATGATTAAGCTGTGGAACCAC 559

Qy         374 GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
                  |||:::|||:::      :::|||   ||||||||
Db         560 AGCCTGCTGGATGCGCGCGCCATGAACACATGCAACACC 598
```

RESULT 9
AV817525/c
LOCUS AV817525 391 bp mRNA linear EST 01-APR-2002
DEFINITION AV817525 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-94-M06 3', mRNA sequence.
ACCESSION AV817525
VERSION AV817525.1 GI:19859415
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 391)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further details.
FEATURES Location/Qualifiers
source 1. .391
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFL09-94-M06"
/clone_lib="RAFL9"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site_1: BamHI; Site_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT 102 a 100 c 73 g 116 t
ORIGIN

Alignment Scores:

Pred. No.:	1.43e-48	Length:	391
Score:	506.00	Matches:	89
Percent Similarity:	87.07%	Conservative:	12
Best Local Similarity:	76.72%	Mismatches:	15
Query Match:	21.28%	Indels:	0
DB:	10	Gaps:	0

US-09-890-220-2 (1-445) x AV817525 (1-391)

```
Qy         274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
               |||   |||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         389 CGAGCTCAGCCCATGGCTCTAGAACAAGTACTTTCGGACCGGGATAGTGAAGATGAAGTT 330

Qy         294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
               |||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db         329 GATGATGATGTGGCAGATTTTGAAGATAGAAGGATGCTCGATGATTTCGTTGATGTGACT 270

Qy         314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
               ||||||||||||||   ||||||:::|||||||||||||||||||||||||||||:::
Db         269 AAAGATGAGAAACAGATGATGCACATGTGGAACTCGTTTGTGAGGAAGCAGCGAGTATTA 210

Qy         334 AlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353
               ||||||||||||||   ||||||||||||||||||||||   :::      :::|||
Db         209 GCAGATGGTCACATTCCATGGGCATGCGAGGCATTCTCAAGATTGCACGGACCCATCATG 150

Qy         354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373
                  |||         |||   ||||||||||||:::|||::::::||||||||||||||||
Db         149 GTTCGAACACCGCACTTGATTTGGTGCTGGAGAGTGTTTATGGTGAAACTGTGGAACCAC 90

Qy         374 GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
               ||||||:::|||:::   |||:::||||||||||||||||   ||||||
Db          89 GGTCTTCTTGATGCCCGAACCATGAACAACTGTAATACCTTTCTCGAA 42
```

RESULT 10
BQ611550
LOCUS BQ611550 569 bp mRNA linear EST 26-JUN-2002
DEFINITION sap61f09.y1 Gm-c1087 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1087-5250 5' similar to TR:Q9ZNT9 Q9ZNT9 FERTILIZATION-INDEPENDENT SEED 2 PROTEIN. ;, mRNA sequence.
ACCESSION BQ611550
VERSION BQ611550.1 GI:21601219
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 569)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

```
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact: ccu@resgen.com web site:
            www.resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 440.
FEATURES              Location/Qualifiers
     source           1. .569
                      /organism="Glycine max"
                      /db_xref="taxon:3847"
                      /clone="SOYBEAN CLONE ID: Gm-c1087-5250"
                      /clone_lib="Gm-c1087"
                      /tissue_type="Roots"
                      /lab_host="DH10B"
                      /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                      XhoI; The mRNA was prepared using polyatract mRNA system
                      from PROMEGA. The cDNA was prepared using the STRATAGENE
                      kit. Complementary DNA was synthesized from mRNA using a
                      primer consisting of a poly(dT) sequence with a XhoI
                      restriction site
                      (5'GAGAGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTT).
                      EcoRI adapters (5'OH-AATTCGGCACGAG and 3'GCCGTGCTCp) were
                      ligated to the blunt-ended cDNA fragments followed by XhoI
                      digestion. The cDNA fragments were directionally cloned
                      into the EcoRI-XhoI restriction sites of the pBluescript
                      vector. The vector was previously dephosphorylated. The
                      ligated cDNA fragments were transformed into DH10B by
                      electroporation. Library is in LB medium with ampicillin
                      and glycerol 8%. Average insert size: 800 bp."
BASE COUNT      172 a    106 c    122 g    169 t
ORIGIN

Alignment Scores:
Pred. No.:                1.9e-47        Length:          569
Score:                    498.50         Matches:         96
Percent Similarity:       82.52%         Conservative:    22
Best Local Similarity:    67.13%         Mismatches:      24
Query Match:              20.96%         Indels:          1
DB:                       14             Gaps:            1

US-09-890-220-2 (1-445) x BQ611550 (1-569)

Qy        233 AlaLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSer 252
              :::|||   |||      |||:::      :::|||||   |||||||:::||||||||
Db         16 AGTAAATCAATATGTAAAAGTGATCCAGCTCTGCCTGCT---AAAACAAAGAAGCTAAGC 72

Qy        253 AlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSer 272
                :::|||||||::::::|||:::   :::|||||||||||||   |||:::||||||
Db         73 ATGGATCGATCAGACTCAAGAAACCGGATGCTTCTGCAGAAGAGACTTTTCTTTCACTCA 132

Qy        273 HisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGlu 292
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        133 CACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGACCGTGATAGTGAAGACGAA 192

Qy        293 ValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspVal 312
              ||||||||||||:::||||||   |||||||||:::|||||||||||||||||||||||
Db        193 GTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTTGATGATTTTGTGGATGTT 252

Qy        313 AsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgVal 332
              :::|||||||||||||||   ||||||||||||||||||:::|||||||||||||||||
Db        253 TCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGGGTG 312

Qy        333 IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGlu 352
              :::||||||||||||:::   |||||||||||||||||||:::   :::   ||||||
Db        313 CTGGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAAGAG 372

Qy        353 LeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
              |||           :::|||||||||         |||   ||||||   :::
Db        373 CTGATCTCATCTCCTGCTTTATTTTGGTACTCCTATCTTCCATTGATATTTATCTAATTT 432

Qy        373 HisGlyLeu 375
                ||||||:::
Db        433 CATGGAGTT 441

RESULT 11
AW234600
LOCUS       AW234600                 474 bp    mRNA    linear   EST 17-JUL-2000
DEFINITION  sf17b08.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1028-208 5' similar to TR:O23525 O23525 HYPOTHETICAL 27.2 KD
            PROTEIN. ;, mRNA sequence.
ACCESSION   AW234600
VERSION     AW234600.1  GI:6566973
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1799   Std Error: 0.00
            High quality sequence stop: 412.
FEATURES              Location/Qualifiers
     source           1. .474
                      /organism="Glycine max"
                      /db_xref="taxon:3847"
                      /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-208"
                      /clone_lib="Gm-c1028"
                      /tissue_type="roots of 'Supernod' plants"
                      /lab_host="DH10B"
                      /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                      XhoI; The mRNA was isolated from roots of Glycine max
                      'Supernod' plants generously donated by Dr. Gary Stacey.
                      The seedlings were innoculated with Bradyrhizobium
                      japonicus, strain USDA110 priot to harvest. Stratagene's
                      cDNA synthesis Kit (catalog number 200401) was used to
                      synthesize the cDNA.  First-strand synthesis was performed
                      with 5-methyl dCTP, hence the ligated cDNA was
                      hemimethylated.  A modification of Stratagene's
                      first-strand synthesis primer was used.  An 'anchor'
                      nucleotide (V=A,C, or G) was added to the 3' end of the
                      primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor
                      the primer at the 5' end of the poly(A) tract.  After
                      second-strand synthesis, the cDNA ends were filled in with
                      cloned Pfu DNA polymerase, ligated to EcoRI adapters and
                      subsequently phosphorylated.  The XhoI site within the
                      first-strand synthesis primer was then restricted by
                      digestion with XhoI; all XhoI sites in the cDNA would be
                      protected by their hemimethylated status.  The cDNA
                      constructs were size-fractionated with a 500bp cutoff,
                      using GibcoBRL Life Technologies' cDNA Size Fractionation
                      column.  The column eluent was then ligated into
                      Stratagene's pBluescript II XR Predigested vector
```

(pBluescript II SK(+) that has been digested with EcoRI and XhoI, and phosporylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 133 a 86 c 107 g 148 t
ORIGIN

Alignment Scores:

Pred. No.:	1.14e-42	Length:	474
Score:	456.50	Matches:	89
Percent Similarity:	78.36%	Conservative:	16
Best Local Similarity:	66.42%	Mismatches:	28
Query Match:	19.20%	Indels:	1
DB:	10	Gaps:	1

US-09-890-220-2 (1-445) x AW234600 (1-474)

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||                :::   ||||||   |||:::   ||||||
Db         43 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 102

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        103 AGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCTCTT 162

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||   |||||||||   |||||||||   |||:::||||||||||
Db        163 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAAAGG 222

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              :::   |||:::   |||:::|||:::||||||||:::|||      |||   |||:::
Db        223 AGGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTTCGG 279

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   |||   ||||||||||||
Db        280 AAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAGCTTT 339

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||:::||||||||   ||||||||||||||||   |||||||||   :::
Db        340 AAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGGGTT 399

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
              |||:::||||||   ||||||||||||||||||||:::|||
Db        400 ACTGAAGATTACCAAGCAGTGAATGTCTCCGTGAAAATTAAT 441
```

RESULT 12
BG580495
LOCUS BG580495 467 bp mRNA linear EST 11-APR-2001
DEFINITION EST482220 GVN Medicago truncatula cDNA clone pGVN-57B2 5' end, mRNA sequence.
ACCESSION BG580495
VERSION BG580495.1 GI:13595559
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 467)
AUTHORS Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town,C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.
TITLE ESTs from one month old nitrogen-fixing root nodules of Medicago truncatula, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M381546e TIGR sequence name: MTCCT01TK More information is available at: http://www.medicago.org
Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES Location/Qualifiers
source 1. .467
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-57B2"
/clone_lib="GVN"
/tissue_type="N2-fixing root nodules"
/dev_stage="effective root nodules harvested one month post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from effective root nodules harvested one month post inoculation with Sinorhizobium meliloti. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XLOLR cells."
BASE COUNT 136 a 87 c 95 g 149 t
ORIGIN

Alignment Scores:

Pred. No.:	6.39e-42	Length:	467
Score:	450.00	Matches:	91
Percent Similarity:	70.32%	Conservative:	18
Best Local Similarity:	58.71%	Mismatches:	40
Query Match:	18.92%	Indels:	6
DB:	12	Gaps:	3

US-09-890-220-2 (1-445) x BG580495 (1-467)

```
Qy         17 SerThrAspGluAsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHis 36
              :::   ||||||:::|||||||||||||||||||||||||      |||||||||   :::
Db          3 ACAGCTGATGAAAGTCTTTTGATTTATTGCAAACCTGTTGAATTCTACAATATTCTCTAT 62

Qy         37 LeuArgSerLeuGlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAla 56
                 |||||||||   ||||||:::||||||   |||||||||   |||||||||   |||
Db         63 CGTCGTTCTCTTCACAATCCTACTTTTCTTAAGAGATGTTTGCGCTATAAAATAAAAGCA 122

Qy         57 LysArgLysArgLysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsn 76
              |||:::||||||:::   |||:::   |||:::|||||||||||||||:::||||||
Db        123 AAGCAAAAAAGGAGGTTGCGAGCA---GGAATTGTGGTTTTCAACTATAGGGACTGTTAC 179

Qy         77 AsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPheCysSerMetLeu 96
              |||   |||:::||||||||||||   ||||||   ||||||||||||||   |||
Db        180 AATGCGCTTCGAAAGACTGAAGTGACTGAAGACTTTTCTTGTCCATTTTGCTTGATGCAG 239

Qy         97 CysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPhe 116
              |||   ||||||||||||||:::|||||||||   :::||||||||||||||   |||
Db        240 TGTGCGAGCTTTAAGGGTTTGCGATTTCATCTTTGTGCATCACATGATCTATTCAACTTT 299

Qy        117 GluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
              ||||||   :::      :::||||||   ||||||||||||||||   :::
Db        300 GAGTTCTGGGTCACCAATGATTACCAAGCAGTGAACGTCTCTGTAAAAACTGAC------ 353

Qy        137 IlePheGluGluGluGlySerAspAsp---------AspLysPheGluProPheSerLeu 153
              |||         |||         |||            :::   :::   |||
Db        354 ATATTGAGATCAGAGAATGTTGCTGATGGAGTAATTCCACAATCACAAACCTTCTTCTTC 413

Qy        154 CysSerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsn 168
              ||||||:::   ||||||||||||   ||||||      |||
Db        414 TGTTCAAGATCTCGAAAGCGTAGACCAAGAGGCTCCGTTCAAAAT 458
```

RESULT 13
BG314044
LOCUS BG314044 405 bp mRNA linear EST 16-APR-2001
DEFINITION WHE2467_D04_G07ZS Triticum monococcum early reproductive apex cDNA

```
            library Triticum monococcum cDNA clone WHE2467_D04_G07, mRNA
            sequence.
ACCESSION   BG314044
VERSION     BG314044.1  GI:13115847
KEYWORDS    EST.
SOURCE      Triticum monococcum.
  ORGANISM  Triticum monococcum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
            ,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
            Stamova,B. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Early reproductive apex cDNA library from Triticum
            monococcum
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Triticum monococcum"
                     /cultivar="DV92"
                     /db_xref="taxon:4568"
                     /clone="WHE2467_D04_G07"
                     /clone_lib="Triticum monococcum early reproductive apex
                     cDNA library"
                     /tissue_type="Early reproductive apex"
                     /dev_stage="Seven week-old plants"
                     /lab_host="E. coli XLOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and
                     poly(A) RNA were prepared from apex at  double-ridge stage
                     to terminal-spikelet stage during transition from
                     vegetative state to flower state, a cDNA library was made,
                     and the  cDNA clones were in vivo excised at the
                     University of California, Davis  (V. Echenique, B. Stamova
                     , J. Dubcovsky).  Plasmid DNA preparations and  DNA
                     sequencing were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT      105 a     85 c    104 g    111 t
ORIGIN

Alignment Scores:
Pred. No.:              8.8e-42         Length:         405
Score:                  448.00          Matches:        78
Percent Similarity:     81.58%          Conservative:   15
Best Local Similarity:  68.42%          Mismatches:     21
Query Match:            18.84%          Indels:         0
DB:                     12              Gaps:           0

US-09-890-220-2 (1-445) x BG314044 (1-405)

Qy        276 GlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAsp 295
              ||||||||||||:::|||||||||:::|||||||   ||||||||||||||||||||||
Db          4 CAGCCAATGGCAGTCGAACAAGTTTTGTCAGATGATGACAGCGAGGATGAAGTTGATGAT 63

Qy        296 AspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAsp 315
              |||:::|||||||   |||||||||:::|||||||   ||||||:::|||||||   ||||||
Db         64 GACATTGCCGACTTGGAAGATAGACGGATGCTTGCTGATTTTCTTGATGTCACGAAAGAT 123

Qy        316 GluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAsp 335
              ||||||      ||||||:::||||||||||||:::||||||||||||||||:::|||||||
Db        124 GAGAAGCTTATTATGCATATGTGGAATTCATTTATTCGGAAACAAAGGGTGCTAGCTGAT 183

Qy        336 GlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArg 355
              |||||||||   ||||||||||||   |||||||||   :::      :::|||  :::
Db        184 GGGCATATACCTTGGGCCTGTGAGGGCTTCTCCCGGCTTCATGGACCGCAGCTTGTACAA 243

Qy        356 TyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeu 375
                  |||   ||||||||||||||||   :::||||||||||||||||||||   |||
Db        244 AACCCTCCTCTGCTCTGGTGCTGGCGTTTGGTGATGATTAAGCTGTGGAACCACAGCCTG 303

Qy        376 ValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
              :::|||:::      :::|||   ||||||||||||||||||
Db        304 CTGGATGCGCGCGCCATGAACACATGCAACACCATTCTTGAG 345

RESULT 14
BQ279630/c
LOCUS       BQ279630                662 bp    mRNA    linear   EST 08-MAY-2002
DEFINITION  1091030A10.x2 1091 - Immature ear with common ESTs screened by
            Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION   BQ279630
VERSION     BQ279630.1  GI:20507433
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 662)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 1091030  row: A  column: 10.
FEATURES             Location/Qualifiers
     source          1. .662
                     /organism="Zea mays"
                     /cultivar="OH43"
                     /db_xref="taxon:4577"
                     /clone_lib="1091 - Immature ear with common ESTs screened
                     by Schmidt lab"
                     /tissue_type="Inflorescence meristem - floral organ
                     primordia"
                     /dev_stage="0.5 cm to 2 cm"
                     /lab_host="Stratagene XLOLR"
                     /note="Organ: Immature ear; Vector: pAD-GAL4; Site_1:
                     EcoRI; Site_2: XhoI; RNA from library 606 was filtered for
                     common ESTs found in 606."
BASE COUNT      208 a    132 c    111 g    211 t
ORIGIN

Alignment Scores:
Pred. No.:              4.76e-41        Length:         662
Score:                  444.50          Matches:        88
Percent Similarity:     64.78%          Conservative:   15
Best Local Similarity:  55.35%          Mismatches:     25
Query Match:            18.69%          Indels:         31
DB:                     14              Gaps:           2

US-09-890-220-2 (1-445) x BQ279630 (1-662)

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
              |||:::         |||:::         ||||||||||||||||||   |||||||||
Db        622 CTTTCACCACCAACAGTACTACAGTTTGGGAAGACAAGGAAGCTATCT---GAGAGATCT 566

Qy        257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
              :::   |||         ||||||||||||||||||||:::||||||||||   |||
Db        565 GACCCTAGAGATCGGCAACTCCTGCAAAAACGACAGTTCTTCCATTCTCACAGGGCGCAG 506
```

```
Qy        277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
              ||||||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        505 CCAATGCAACTGGAGCAAGTGTTCTCGGACCGTGATAGTGAAGATGAAGTTGATGATGAT 446

Qy        297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
              :::||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        445 ATTGCTGACTTCGAGGATAGAAGAATGCTTGATGATTTTGTTGATGTTACGAAAGATGAA 386

Qy        317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
              |||      ||||||:::||||||||||||||||||||||||
Db        385 AAACTTATTATGCATATGTGGAATTCATTTGTTCGAAAACAAAGG---------------- 341

Qy        337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356

Db        341 ------------------------------------------------------------ 341

Qy        357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
                          ||||||   |||:::|||||||||||||||||   ::::::
Db        340 ---------------GGTTGGCGTTTCTTCATGATTAAACTTTGGAACCATAACATTTTA 296

Qy        377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
              |||:::   |||:::|||   |||||||||:::|||:::      :::      |||
Db        295 GATGCCCGCACTATGAACACATGCAATACAGTCCTTCAAATTTTACAAGAAGAAAGC 239

RESULT 15
BE203328
LOCUS       BE203328                612 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  EST403350 KV1 Medicago truncatula cDNA clone pKV1-5M13, mRNA
            sequence.
ACCESSION   BE203328
VERSION     BE203328.1  GI:8746599
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   VandenBosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
            Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
            Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula 24 hours after inoculation
            with Sinorhizobium meliloti
  JOURNAL   Unpublished (1999)
COMMENT     Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@cbs.umn.edu
            Texas A&M University:T263100e
            TIGR sequence name:MTIAM79TK
            More information is available at. .
            http://chrysie.tamu.edu/medicago
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .612
                     /organism="Medicago truncatula"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pKV1-5M13"
                     /clone_lib="KV1"
                     /tissue_type="Seedling roots"
                     /dev_stage="24 hours post-inoculation with Sinorhizobium
                     meliloti"
                     /lab_host="E. coli strain XLOLR"
                     /note="Vector: pBluescript SK -; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts.  Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      200 a     88 c    132 g    192 t
ORIGIN
```

Alignment Scores:

Pred. No.:	7.23e-41	Length:	612
Score:	442.50	Matches:	79
Percent Similarity:	81.15%	Conservative:	20
Best Local Similarity:	64.75%	Mismatches:	18
Query Match:	18.61%	Indels:	5
DB:	10	Gaps:	1

US-09-890-220-2 (1-445) x BE203328 (1-612)

```
Qy        293 ValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspVal 312
              |||||||||   ||||||:::   |||||||||:::|||||||||||||||||||||||
Db          3 GTTGATGATGGAGTTGCTAACCTGGAAGATAGAAAGATGCTTGATGATTTTGTGGATGTT 62

Qy        313 AsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgVal 332
              :::|||||||||||||||   |||||||||||||||||||||:::|||||||||||||||
Db         63 TCCAAAGATGAAAAGCAGCTCATGCATCTCTGGAACTCTTTTATGAGAAAACAAAGGGTA 122

Qy        333 IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGlu 352
              :::|||||||||||||:::   |||||||||||||||||||||||:::   |||   ||||||
Db        123 CTGGCTGATGGTCATGTTCCGTGGGCTTGTGAGGCATTTTCCAAGCATTATGCAAAAGAG 182

Qy        353 LeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
              |||:::         :::|||||||||||||||||||||||||:::::::||||||||||||
Db        183 CTGAATTCATCCCGAACTTTATTTTGGTGTTGGAGATTATTCATGGTCAAACTTTGGAAT 242

Qy        373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
              |||||||||:::|||:::::::|||:::||||||||||:::   |||||||||:::   :::
Db        243 CATGGGCTTCTTGATGCATCAACAATGAACAACTGCAGTATAATATTGGAAAGCAACCAA 302

Qy        393 AsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSer 412
              |||            :::   ||||||                  |||||||||   |||
Db        303 AACGGGGGATCAGACAGTGGCACAAAT--------------TGAAGACCATCTGTCTCA 347

Qy        413 AsnThr 414
              :::|||
Db        348 TCCACA 353
```

Search completed: June 20, 2003, 01:23:27
Job time : 1937 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:23:13 ; Search time 23 Seconds
(without alignments)
802.477 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	138.5	5.8	989	1	PTP3_DICDI	P54637	dictyosteli
2	132	5.6	749	1	MAD1_YEAST	P40957	saccharomyc
3	127.5	5.4	448	1	AAC2_DICDI	P14196	dictyosteli
4	126.5	5.3	1584	1	KYK1_DICDI	P18160	dictyosteli
5	123.5	5.2	1858	1	P3K2_DICDI	P54674	dictyosteli
6	120	5.0	1000	1	S155_YEAST	P43612	saccharomyc
7	120	5.0	1195	1	YK76_YEAST	P36168	saccharomyc
8	119	5.0	1094	1	YB00_YEAST	P38114	saccharomyc
9	117.5	4.9	988	1	OMB_DROME	Q24432	drosophila
10	117	4.9	2175	1	HMCU_DROME	P10180	drosophila
11	114	4.8	964	1	YIN0_YEAST	P40467	saccharomyc
12	112.5	4.7	2278	1	FAB1_YEAST	P34756	saccharomyc
13	112	4.7	1314	1	SWI1_YEAST	P09547	saccharomyc
14	111.5	4.7	735	1	CIGB_DICDI	Q94481	dictyosteli
15	111.5	4.7	858	1	CYAG_DICDI	Q03101	dictyosteli
16	108	4.5	1407	1	CYAA_DICDI	Q03100	dictyosteli
17	107.5	4.5	833	1	HSF_YEAST	P10961	saccharomyc
18	107.5	4.5	2339	1	RPC1_PLAFA	P27625	plasmodium
19	106.5	4.5	1585	1	P3K3_DICDI	P54675	dictyosteli
20	106	4.5	694	1	HAA1_YEAST	Q12753	saccharomyc
21	106	4.5	779	1	SRP_DROME	P52172	drosophila
22	105.5	4.4	558	1	YCXA_ASTLO	P34782	astasia lon
23	105	4.4	396	1	RMAR_YEAST	P02381	saccharomyc
24	105	4.4	555	1	POST_XENLA	P27536	xenopus lae
25	105	4.4	732	1	KMHB_DICDI	P90648	dictyosteli
26	104.5	4.4	666	1	YEA7_YEAST	P40002	saccharomyc
27	104.5	4.4	935	1	COPG_YEAST	P32074	saccharomyc
28	104.5	4.4	1419	1	MDR_PLAFF	P13568	plasmodium
29	104	4.4	871	1	SC10_YEAST	Q06245	saccharomyc
30	103	4.3	432	1	AP2_ARATH	P47927	arabidopsis
31	103	4.3	490	1	CAR3_DICDI	P35352	dictyosteli
32	103	4.3	1085	1	CUT7_SCHPO	P24339	schizosacch
33	102.5	4.3	481	1	BIND_STRPU	P06651	strongyloce
34	102.5	4.3	1076	1	RPOB_ASTLO	P27059	astasia lon
35	102.5	4.3	1570	1	P3K1_DICDI	P54673	dictyosteli
36	102.5	4.3	2733	1	RRPB_CVMA5	P16342	murine coro
37	101.5	4.3	647	1	KNRL_DROME	P13054	drosophila
38	101.5	4.3	946	1	FTSK_CAMJE	Q46089	campylobact
39	101.5	4.3	1341	1	YL78_YEAST	Q05854	saccharomyc
40	101.5	4.3	2731	1	RRPB_CVMJH	P29982	murine coro
41	101	4.2	534	1	GCR2_YEAST	Q01722	saccharomyc
42	101	4.2	537	1	ARP_PLAFA	P04931	plasmodium
43	101	4.2	1131	1	YAB9_YEAST	P31380	saccharomyc
44	100.5	4.2	1188	1	OSH1_YEAST	P35845	saccharomyc
45	100	4.2	284	1	SMX5_SCHMA	Q26604	schistosoma

ALIGNMENTS

RESULT 1
PTP3_DICDI
ID PTP3_DICDI STANDARD; PRT; 989 AA.
AC P54637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (Protein-tyrosine-
DE phosphate phosphohydrolase 3).
GN (PTPC1 OR PTP3) AND (PTPC2 OR PTP3).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96189126; PubMed=8628311;
RA Gamper M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during
RT Dictyostelium growth and development.";
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -!- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 kDa (P130).
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH
CC AND DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC ---

CC ---
DR EMBL; U38197; AAC47041.1; -.
DR HSSP; Q06124; 2SHP.
DR DictyDb; DD01111; ptpC1.
DR DictyDb; DD0????; ptpC2.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT ACT_SITE 649 649 BY SIMILARITY.
FT DOMAIN 460 716 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 64 71 POLY-ASN.
FT DOMAIN 109 118 POLY-ASN.

```
FT   DOMAIN       137    190       POLY-ASN.
FT   DOMAIN       249    257       POLY-SER.
FT   DOMAIN       258    265       POLY-THR.
FT   DOMAIN       286    289       POLY-ASN.
FT   DOMAIN       366    371       POLY-SER.
FT   DOMAIN       787    790       POLY-GLN.
FT   DOMAIN       834    839       POLY-GLN.
FT   DOMAIN       883    892       POLY-GLN.
FT   DOMAIN       906    914       POLY-ASN.
FT   DOMAIN       943    963       POLY-ASN.
SQ   SEQUENCE   989 AA;  109995 MW;  9371105AF80974AF CRC64;

  Query Match             5.8%;  Score 138.5;  DB 1;  Length 989;
  Best Local Similarity   19.8%;  Pred. No. 0.0064;
  Matches   72;  Conservative   51;  Mismatches  140;  Indels  101;  Gaps   12;

Qy        129 VSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGT 188
              :|: :|    |  |:|  :       |    |||| |   |          ||
Db        660 ISININMAKIERFGNDPSQMNISIKDSVLELRRQRRGMVQT---------LDQYIFIFKV 710

Qy        189 ENGITLLNDGNRGLGYPEATELAGQFEMTS------NIPPAIA------HSSLDAGAKVI 236
              | : | : | | |  |         : |       :||| :        ||:     :|
Db        711 INDV-LTDMGIRSLSSPSKRRSCEMIKSTPMPRLDISIPPPLTFTPKDFQSSISPSTDMI 769

Qy        237 LTSEAVVPATKTRKLSAERSE----ARSHLLL-------------QKRQFYHSHRVQP-M 278
              :   :   |:| |   :: :    ::| : :             : :|| |  ::|| :
Db        770 ASLSIITQMTQTLKFPPQQQQDNPFSKSSIKISPSPLNSTNISIPKNQQFQHPFQIQPQL 829

Qy        279 ALEQVMSDRDSEDEVDDD--------------VADFEDRQMLDDFVDVNKDEKQFMHLWN 324
              |     : |  :::|:              |       : :|  : : : ||
Db        830 DLNLQQQQQQSSQQLNDNPPLNMSSNSIKFPPVTSLSSCHLFEDSKNNDNNNKQ------ 883

Qy        325 SFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNC 384
                ::|:       :  |  || |                                :||
Db        884 ---QQQQQQQQQKNNQQCSGFSHF-------------------------------LNNN 908

Qy        385 NTILENCR-----NSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNN- 438
              |    |       | |    :||: |    |: : || |  :: |: || ||  | |||
Db        909 NNNDNNGSSGGGFNGSFLFNSNNSGSSSTNSECSNNNKNNNNNSNNNNNNNNNKNSDNNG 968

Qy        439 SRDK 442
              ::||
Db        969 TKDK 972


RESULT 2
MAD1_YEAST
ID   MAD1_YEAST     STANDARD;      PRT;   749 AA.
AC   P40957;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Spindle assembly checkpoint component MAD1 (Mitotic MAD1 protein).
GN   MAD1 OR YGL086W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W303;
RX   MEDLINE=96042315; PubMed=7593191;
RA   Hardwick K.G., Murray A.W.;
RT   "Madlp, a phosphoprotein component of the spindle assembly checkpoint
RT   in budding yeast.";
RL   J. Cell Biol. 131:709-720(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=97435481; PubMed=9290212;
RA   Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT   "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT   chromosome VII.";
RL   Yeast 13:1077-1090(1997).
RN   [3]
RP   INTERACTIONS.
RX   MEDLINE=98128031; PubMed=9461437;
RA   Hwang L.H., Lau L.F., Smith D.L., Mistrot C.A., Hardwick K.G.,
RA   Hwang E.S., Amon A., Murray A.W.;
RT   "Budding yeast Cdc20: a target of the spindle checkpoint.";
RL   Science 279:1041-1044(1998).
CC   -!- FUNCTION: CENTRAL COMPONENT OF THE SPINDLE ASSEMBLY CHECKPOINT.
CC   -!- SUBUNIT: THE SPINDLE CHECKPOINT COMPLEX IS COMPOSED OF MAD1, MAD2
CC       AND MAD3. IT INTERACTS WITH CDC20.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- PTM: BECOMES HYPERPHOSPHORYLATED WHEN WILD-TYPE CELLS ARE ARRESTED
CC       IN MITOSIS.
CC   -!- SIMILARITY: SOME, TO S.POMBE SPBC3D5.04C.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U14632; AAA91620.1; -.
DR   EMBL; Z72608; CAA96791.1; -.
DR   PIR; S48246; S48246.
DR   SGD; S0003054; MAD1.
KW   Mitosis; Coiled coil; Nuclear protein; Phosphorylation.
FT   DOMAIN       57    221       COILED COIL (POTENTIAL).
FT   DOMAIN      253    324       COILED COIL (POTENTIAL).
FT   DOMAIN      390    656       COILED COIL (POTENTIAL).
FT   DOMAIN      323    329       POLY-ASP.
FT   DOMAIN      330    347       POLY-ASN.
FT   DOMAIN      355    363       POLY-ASN.
FT   DOMAIN      601    604       POLY-LEU.
SQ   SEQUENCE   749 AA;  87651 MW;  312DDE1BE241610D CRC64;

  Query Match             5.6%;  Score 132;  DB 1;  Length 749;
  Best Local Similarity   21.3%;  Pred. No. 0.014;
  Matches   87;  Conservative   57;  Mismatches  143;  Indels  122;  Gaps   18;

Qy         89 SCPFCSMLCGS-------------FKGLQFHLNSSHDLFEFEEFKLFEEYQTVNVSVKLNS 135
              | ||     ||               : ||| ||:  : :| | ||  : ||  :  | :
Db         40 SSPFLESPGGSPDVGSTNGQSNRQIQALQFKLNTLQNEYEIE-KLQLQKQTNILEKKYKA 98

Qy        136 FIFE-EEGSDDDKF-----EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTE 189
               | | |:  :| |:      :       | | |      |:  :   |   | :
Db         99 TIDELEKALNDTKYLYESNDKLEQELKSLKERSANSMNDKDK---CIEELRT-------- 147

Qy        190 NGITLLNDGNRGLGYPEATELAGQFE-MTSNIPPAIAHSSLDAGAKVILTSEAVVPATKT 248
                 || |          |   |  |::   | :     |  |:| :
Db        148 ---TLQNKD------LEMETLRQQYDSKLSKVTNQCDHFKLEAES--------------- 183

Qy        249 RKLSAERSEARSHLLLQ------KRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFED 302
                        || ||       |||      :|     |||       | ||::: :
Db        184 ----------SHSLLMKYEKEIKRQSVDIKDLQ----HQVM-------EKDDELSSVKA 221

Qy        303 RQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIAD----------GHISWACEAFSRFYEKE 352
               :|::   : : :|     :|     ::| |                ::  | |     |:
Db        222 SKMINSHPNYSTEE------FNELTEMNKMIQDQVQYTKELELANMQQANEL------KK 269

Qy        353 LHRYSSLFWCWRLFLIKLWN-----HGL---VDSATINNCN--------TILENCRNSSD 396
              | :     |:|   || |     | |     ::   : | :          |  :  :  |
Db        270 LKQSQDTSTFWKLENEKLQNKLSQLHVLESQYENLQLENIDLKSKLTKWEIYNDSDDDDD 329

Qy        397 TTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
                 ||:|: :  :|:| :|||   : |:||| |   |   |: :::|:
Db        330 NNVNNNDNNNNNKNDNNNDNNNDTSNNNNINNNNRTKNNIRNNPEEIIR 378


RESULT 3
AAC2_DICDI
```

```
ID   AAC2_DICDI     STANDARD;      PRT;   448 AA.
AC   P14196;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   AAC-rich mRNA clone AAC11 protein (Fragment).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90066348; PubMed=2511421;
RA   Shaw D.R., Richter H., Giorda R., Ohmachi T., Ennis H.L.;
RT   "Nucleotide sequences of Dictyostelium discoideum developmentally
RT   regulated cDNAs rich in (AAC) imply proteins that contain clusters of
RT   asparagine, glutamine, or threonine.";
RL   Mol. Gen. Genet. 218:453-459(1989).
CC   -!- DEVELOPMENTAL STAGE: THE CONCENTRATION OF AAC-RICH MRNAS IS LOW
CC       IN DORMANT SPORES AND GROWING CELLS, BUT INCREASES DURING
CC       SPORE-GERMINATION AND MULTICELLULAR DEVELOPMENT.
CC   -!- MISCELLANEOUS: SEVERAL PROTEINS DERIVE FROM AAC-RICH MRNA, WHICH,
CC       DUE TO A FRAMESHIFT ALSO HAVE ACA AND CAA CODONS AND THUS ARE
CC       ASN-, THR- OR GLN-RICH.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16522; CAA34529.1; -.
DR   PIR; S05355; S05355.
DR   DictyDb; DD05006; -.
DR   InterPro; IPR000637; AT_hook.
DR   InterPro; IPR000116; Highmoblty_IY.
DR   Pfam; PF02178; AT_hook; 4.
DR   ProDom; PD005593; Highmoblty_IY; 1.
DR   SMART; SM00384; AT_hook; 4.
KW   Repeat.
FT   NON_TER       1      1
FT   DOMAIN        5     25       GLN-RICH.
FT   DOMAIN      115    144       ASN-RICH.
FT   DOMAIN      268    384       ASN-RICH.
SQ   SEQUENCE   448 AA;  48636 MW;  D8FD80D910D99817 CRC64;

  Query Match             5.4%;  Score 127.5;  DB 1;  Length 448;
  Best Local Similarity   35.7%;  Pred. No. 0.016;
  Matches   25;  Conservative   20;  Mismatches   24;  Indels    1;  Gaps    1;

Qy        372 NHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNN 431
              |:   ::: ||| |     |  |::: :: ||||: :  :::|||||   : |  || ||
Db        273 NNNNNNNSNINNINNNNNNNSNNNNNSSNNNNNN-NNSTNNNTNNNNNNTNNNTNNNNNN 331

Qy        432 VDNKDNNSRD 441
              ::| :||: :
Db        332 INNNNNNTNN 341


RESULT 4
KYK1_DICDI
ID   KYK1_DICDI     STANDARD;      PRT;  1584 AA.
AC   P18160;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
DE   protein kinase 1).
GN   PYKA OR SPLA OR DPYK1.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JH10;
RX   MEDLINE=97053827; PubMed=8898241;
RA   Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;
RT   "The Dictyostelium dual-specificity kinase splA is essential for
RT   spore differentiation.";
RL   Development 122:3295-3305(1996).
RN   [2]
RP   SEQUENCE OF 1248-1584 FROM N.A.
RX   MEDLINE=90287147; PubMed=1972546;
RA   Tan J.L., Spudich J.A.;
RT   "Developmentally regulated protein-tyrosine kinase genes in
RT   Dictyostelium discoideum.";
RL   Mol. Cell. Biol. 10:3578-3583(1990).
CC   -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC       DURING THE MOUND STAGE OF MORPHOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC   -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U32174; AAB41125.1; -.
DR   EMBL; M33785; AAA33202.1; -.
DR   PIR; A35670; A35670.
DR   DictyDb; DD03010; pykA.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR001660; SAM.
DR   InterPro; IPR003878; SPRY_domain.
DR   InterPro; IPR003877; SPRY_receptor.
DR   InterPro; IPR004040; STY_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; pkinase; 1.
DR   Pfam; PF00536; SAM; 1.
DR   Pfam; PF00622; SPRY; 3.
DR   ProDom; PD000001; Euk_pkinase; 1.
DR   SMART; SM00454; SAM; 1.
DR   SMART; SM00449; SPRY; 3.
DR   SMART; SM00221; STYKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT   DOMAIN      908    972       SAM.
FT   DOMAIN      403    420       POLY-ASN.
FT   DOMAIN      428    435       POLY-THR.
FT   DOMAIN      449    480       POLY-ASN.
FT   DOMAIN      483    491       POLY-ASN.
FT   DOMAIN      494    508       POLY-ASN.
FT   DOMAIN      512    532       POLY-ASN.
FT   DOMAIN      596    600       POLY-ASN.
FT   DOMAIN      808    811       POLY-PHE.
FT   DOMAIN     1026   1029       POLY-SER.
FT   DOMAIN     1195   1210       POLY-ASN.
FT   DOMAIN     1215   1220       POLY-GLN.
FT   DOMAIN     1224   1233       POLY-GLN.
FT   DOMAIN     1266   1274       POLY-PRO.
FT   DOMAIN     1289   1561       PROTEIN KINASE.
FT   NP_BIND    1295   1303       ATP (BY SIMILARITY).
FT   BINDING    1316   1316       ATP (BY SIMILARITY).
FT   ACT_SITE   1417   1417       BY SIMILARITY.
FT   CONFLICT   1248   1248       D -> R (IN REF. 2).
FT   CONFLICT   1435   1435       V -> L (IN REF. 2).
```

```
SQ   SEQUENCE   1584 AA;  174304 MW;  5D1589458D8E01E3 CRC64;

  Query Match             5.3%;  Score 126.5;  DB 1;  Length 1584;
  Best Local Similarity   22.2%;  Pred. No. 0.099;
  Matches   77;  Conservative   37;  Mismatches  100;  Indels  133;  Gaps   16;

Qy        149 EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEAT 208
              | | :|    |:    |::   | :  |||||||  | |   |:        | |
Db        210 EEFEVCRWSEKKNYH-GKHVVVRNRTAFLPLDSPKDTIG---GVRATQPFGEGFCY---- 261

Qy        209 ELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQ 268
                  ||        :    || |   |  :    |                |:    |
Db        262 -----FE--------VIIDQLDKGQLSIGLANLEYPTF------------YHVGWMPRS 295

Qy        269 F-YHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
              . : ||                        :||   |  |:       ||: |      : |
Db        296 YGYH----------------------NDDGRKFRWREE----PGVNEGES-----YGSSY 324

Qy        328 RKQRVIADGHISWACE-------------AFSRFYEKELHRYSSLFWCWRLF-------- 366
              :|  :|   | :|:              |||  |        :|:    |
Db        325 KKGDIIGCG-LSFTSREIFFTKNGMYLGTAFSNVY--------GVFYPSVAFNEPGISIT 375

Qy        367 ------------LIKLWNHGLVDSATI----------------NNCNTILENCRNSSDTT 398
                         : | |   |:| :|                 || | |: | : :: ||
Db        376 GVFGPPFKFSQVTLMLKN---VNSTSILVPNGNNNNNSNNNNNNNNNIIGNGKITTTTT 432

Qy        399 TT------NNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNS 439
              |:      |||  :   :::| ||||  :: |: || || :| :|||
Db        433 TSTSPSSINNNEDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNS 479


RESULT 5
P3K2_DICDI
ID   P3K2_DICDI     STANDARD;      PRT;  1858 AA.
AC   P54674;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Phosphatidylinositol 3-kinase 2 (EC 2.7.1.137) (PI3-kinase)
DE   (PtdIns-3-kinase) (PI3K).
GN   PIKB OR PIK2.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RX   MEDLINE=96009592; PubMed=7565716;
RA   Zhou K., Takegawa K., Emr S.D., Firtel R.A.;
RT   "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
RT   discoideum: biological roles of putative mammalian p110 and yeast
RT   Vps34p PI 3-kinase homologs during growth and development.";
RL   Mol. Cell. Biol. 15:5645-5656(1995).
CC   -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
CC       1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC   -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23477; AAA85722.1; -.
DR   DictyDb; DD01100; pikB.
DR   InterPro; IPR002420; PI3K_C2.
DR   InterPro; IPR000341; PI3K_ras_bind.
DR   InterPro; IPR001263; PI3Ka.
DR   InterPro; IPR000403; PI3_PI4_kinase.
DR   Pfam; PF00454; PI3_PI4_kinase; 1.
DR   Pfam; PF00613; PI3Ka; 1.
DR   Pfam; PF00792; PI3K_C2; 1.
DR   Pfam; PF00794; PI3K_rbd; 1.
DR   SMART; SM00142; PI3K_C2; 1.
DR   SMART; SM00144; PI3K_rbd; 1.
DR   SMART; SM00145; PI3Ka; 1.
DR   SMART; SM00146; PI3Kc; 1.
DR   PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR   PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR   PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW   Transferase; Kinase; Multigene family.
FT   DOMAIN       34     40       POLY-GLY.
FT   DOMAIN      166    172       POLY-SER.
FT   DOMAIN      185    226       POLY-ASN.
FT   DOMAIN      227    235       POLY-THR.
FT   DOMAIN      246    253       POLY-SER.
FT   DOMAIN      259    268       POLY-ASN.
FT   DOMAIN      294    303       POLY-THR.
FT   DOMAIN      361    364       POLY-SER.
FT   DOMAIN      369    384       POLY-SER.
FT   DOMAIN      425    429       POLY-GLN.
FT   DOMAIN      439    444       POLY-THR.
FT   DOMAIN      445    454       POLY-SER.
FT   DOMAIN      562    570       POLY-GLY.
FT   DOMAIN      715    727       POLY-THR.
FT   DOMAIN      982    990       POLY-GLN.
FT   DOMAIN     1015   1049       POLY-ASN.
FT   DOMAIN     1598   1858       PI3K/PI4K.
SQ   SEQUENCE   1858 AA;  203945 MW;  A6C0333D4CDDEA03 CRC64;

  Query Match             5.2%;  Score 123.5;  DB 1;  Length 1858;
  Best Local Similarity   27.5%;  Pred. No. 0.21;
  Matches   38;  Conservative   25;  Mismatches   42;  Indels   33;  Gaps    5;

Qy        312 VNKDEKQFMH-LWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKL 370
              : |::|: :: : || | |          | |          |:|  |
Db        113 IEKEKKEIINTISNSGVTKYMT--------ALEILDSTINTPLNRSRS------------ 152

Qy        371 WNHGLVDSATINNCNTILENCRNS-------SDTTTTNNNNSVDRPSDSNTNNNNIVDHP 423
                  | : |  |  || :  :  :|       | |||:||||: : :::| ||||  ::
Db        153 ---GSIGSKPI--CNNLTSSSSSSSTTATTPSPTTTSNNNNNNNNNNNNNNNNNNNNNN 207

Qy        424 NDINNKNNVDNKDNNSRD 441
              |: || || :| :||: :
Db        208 NNNNNNNNNNNNNNNNN 225


RESULT 6
S155_YEAST
ID   S155_YEAST     STANDARD;      PRT;  1000 AA.
AC   P43612;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   SIT4-associating protein SAP155.
GN   SAP155 OR YFR040W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=96220458; PubMed=8649382;
RA   Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA   Arndt K.T.;
RT   "The SAP, a new family of proteins, associate and function positively
RT   with the SIT4 phosphatase.";
RL   Mol. Cell. Biol. 16:2744-2755(1996).
RN   [2]
RP   SEQUENCE OF 98-1000 FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=95400292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
```

```
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
RN   [3]
RP   SEQUENCE OF 98-1000 FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=96287652; PubMed=8686379;
RA   Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA   Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT   "Analysis of a 36.2 kb DNA sequence including the right telomere of
RT   chromosome VI from Saccharomyces cerevisiae.";
RL   Yeast 12:149-167(1996).
CC   -!- FUNCTION: ASSOCIATES WITH THE SIT4 PHOSPHATASE IN A CELL CYCLE
CC       DEPENDENT MANNER. MAY BE DIRECTLY OR INDIRECTLY INVOLVED IN
CC       SIT4-DEPENDENT FUNCTIONS IN BUDDING AND IN NORMAL G1 CYCLIN
CC       EXPRESSION.
CC   -!- PTM: HYPERPHOSPHORYLATED IN THE ABSENCE OF SIT-4.
CC   -!- SIMILARITY: BELONGS TO THE SAPS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U50560; AAC49303.1; -.
DR   EMBL; D50617; BAA09279.1; -.
DR   SGD; S0001936; SAP155.
KW   Phosphorylation; Cell cycle.
FT   CONFLICT    663    663       N -> T (IN REF. 2 AND 3).
FT   CONFLICT    674    689       TYSKSNYMIRDCFQNN -> DLFKIKLYDTRIVSKIM (IN
FT                                REF. 2 AND 3).
FT   CONFLICT    814    815       DY -> EL (IN REF. 2 AND 3).
FT   CONFLICT    818    822       RYSSN -> PDIQVI (IN REF. 2 AND 3).
SQ   SEQUENCE   1000 AA;  114880 MW;  AD4FBF0BC7C588D5 CRC64;

  Query Match             5.0%;  Score 120;  DB 1;  Length 1000;
  Best Local Similarity   24.0%;  Pred. No. 0.17;
  Matches   61;  Conservative   24;  Mismatches   65;  Indels  104;  Gaps   12;

Qy        282 QVMSDRDSEDEVDDDVADFE--------DRQMLDDFV-DVNKDEKQFMHLWNSFVRKQRV 332
              | |:|    ||   ||: |          | ::  ||:    :|  :|  |     ||  :
Db        737 QFMTDIVISDEKGTDVSRFSPVIRDPNFDFKITTDFILRGYQDSYKFYEL-----RKMNL 791

Qy        333 IADGHISWACE---AFSRFYEKE-LHRYSSLF----WCW---------RLFLIKLWNHG- 374
                |||    |    ||: |: : : ||||      | :         |:   |:   |
Db        792 GYMGHIVLIAEEVVKFSKLYKVDYISRYSSNLQTEEWQYYSEEVLNETRMMYSKILGGGS 851

Qy        375 -----------LVDSATI-------NNCNTILENCRNSS-------------------- 395
                        | |: |:       :| | ||::   ||
Db        852 YIDDGNGNIIPQLPDNTTVLTPNGDASNNNEILDSDTGSSNGTSGGGQLINVESLEEQLS 911

Qy        396 --------------------DTTTTNNNNSV-------DRPSDSNTNNNNIVDHPNDIN 427
                               |    |  | |        |: |:|| || |     :: |
Db        912 LSTESDLHNKLREMLINRAQEDVDNKNTENGVFILGPPEDKNSNSNINNTN-----HNSN 966

Qy        428 NKNNVDNKDNNSRD 441
              | || || |||  |
Db        967 NSNNNDNNDNNDND 980


RESULT 7
YK76_YEAST
ID   YK76_YEAST     STANDARD;      PRT;  1195 AA.
AC   P36168;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Hypothetical 137.5 kDa protein in MPL1-PPC1 intergenic region.
GN   YKR096W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE OF 1-843 FROM N.A.
RA   Baladron V., Ballesta J.P.G., Bou G., del Rey F., Esteban P.F.,
RA   Garcia-Cantalejo J.M., Garcia-Ramirez J.J., Gonzalez A., Jimenez A.,
RA   Revuelta J.L., Santos M.A.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 737-1195 FROM N.A.
RA   Gaillon L., Dujon B.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: STRONG, TO YEAST YIL151C. ALSO TO THE N-TERMINAL OF
CC       YEAST ALDEHYDE DEHYDROGENASE, MITOCHONDRIAL 1 (ALD1).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z28321; CAA82176.1; -.
DR   PIR; S38174; S38174.
DR   SGD; S0001804; YKR096W.
KW   Hypothetical protein; ATP-binding.
FT   DOMAIN      210    250       ASN-RICH.
FT   NP_BIND     465    472       ATP (POTENTIAL).
SQ   SEQUENCE   1195 AA;  137490 MW;  D7251DED6B523622 CRC64;

  Query Match             5.0%;  Score 120;  DB 1;  Length 1195;
  Best Local Similarity   21.4%;  Pred. No. 0.22;
  Matches   62;  Conservative   37;  Mismatches   99;  Indels   92;  Gaps   10;

Qy        166 RNNTRRLKVCFLPLDSPSLTNGTENGI-TLLNDGNRGLGYPEATELAGQ---------- 213
              || |  | :    : :||      | :   :|   | :  |  ::: |:
Db         42 RNTTANLSLSDFQVLNPSSKRQNSNSVYDDINSSKRRISRPRFSDIEGKNNDHTYPERTT 101

Qy        214 -FEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKL-SAERSEARSHLLLQKRQFYH 271
               |   | |    ||     |  |  |  |  |::  | |   ||    :  :|  :
Db        102 VKESEKNPSPRYVSSS-----KRALKRENSVGITQSSALISKSFSENGGSIAHEK--WSP 154

Qy        272 SHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQR 331
              : ::|: : |            : :|          |||   ||:    :   | ||
Db        155 ENMIKPLNVSQ-----------NSLA----------FVDAGSDEQSKSEIVGGFQRKSN 192

Qy        332 VIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENC 391
                                                        :|  ||: :   :
Db        193 ----------------------------------------------NSQEINDKD---NSA 204

Qy        392 RNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRD 441
              |:     : ||||: :  |::| ||||  |  |: || |: || :||| |
Db        205 RDQDFNNSGNNNNNNNHSSNNNDNNNNNNDDNNNNNNSNSRDN-NNNSDD 253


RESULT 8
YB00_YEAST
ID   YB00_YEAST     STANDARD;      PRT;  1094 AA.
AC   P38114;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative 126.9 kDa transcriptional regulatory protein in YSW1-RIB7
DE   intergenic region.
GN   YBR150C OR YBR1133.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
RA   Niegemann E., Schenk-Groeninger R., Servos J., Wehner E.,
RA   Wolter R., Brendel M., Bauer J., Braun H., Dern K., Duesterhus S.,
RA   Gruenbein R., Hedges D., Kiesau P., Korol S., Krems B., Proft M.,
RA   Siegers K., Baur A., Boles E., Miosga T.,
RA   Schaaff-Gerstenschlaeger I., Zimmermann F.K.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC       CLUSTER DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z36019; CAA85108.1; -.
DR   PIR; S46021; S46021.
DR   HSSP; P08657; 1CLD.
DR   TRANSFAC; T03677; -.
DR   SGD; S0000354; YBR150C.
DR   InterPro; IPR001138; Fungi_TrN.
DR   Pfam; PF00172; Zn_clus; 1.
DR   SMART; SM00066; GAL4; 1.
DR   PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR   PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW   Hypothetical protein; Transcription regulation; DNA-binding;
KW   Nuclear protein; Zinc; Metal-binding.
FT   DNA_BIND     107     137       ZN(2)-CYS(6), FUNGAL-TYPE.
SQ   SEQUENCE   1094 AA;  126902 MW;  3C08F83F5879D14F CRC64;

  Query Match             5.0%;  Score 119;  DB 1;  Length 1094;
  Best Local Similarity   27.5%;  Pred. No. 0.23;
  Matches   44;  Conservative   26;  Mismatches   44;  Indels   46;  Gaps    9;

Qy        303 RQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACE-------AFSRFYE---KE 352
              |:  || |:: |          ||: |::: | |  |:  |       :|: | |   ::
Db        859 RESKDDMVELLK---------TSFINKEKMAAFG--SFVTEDQEKMEVSFNIFNEITIQD 907

Qy        353 LH--RYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNS-VDRP 409
              |:  ::||       : ||| : ::          |   :|: | | ||  :  |
Db        908 LNFLQFSS--------IPKLWENKTLEPG---------EEYHHSNGTNTDNNETTGADDT 950

Qy        410 SDSNTNNNNIVDHPNDI-----NNKNNVDNKDNNSRDKVI 444
                |:| ||||   : |:      || ||  | :||  |  |
Db        951 DDNNNNNNNNNKNGNNSSSTINNNNNNYSNSNNNDNDNNI 990


RESULT 9
OMB_DROME
ID   OMB_DROME      STANDARD;      PRT;   988 AA.
AC   Q24432; Q27917; Q9W4K5;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Optomotor-blind protein (Lethal(1)optomotor-blind) (L(1)omb) (Bifid
DE   protein).
GN   BI OR OMB OR CG3578.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oregon-R; TISSUE=Embryo;
RX   MEDLINE=92159016; PubMed=1741374;
RA   Pflugfelder G.O., Roth H., Poeck B., Kerscher S., Schwarz H.,
RA   Jonschker B., Heisenberg M.;
RT   "The lethal(1)optomotor-blind gene of Drosophila melanogaster is a
RT   major organizer of optic lobe development: isolation and
RT   characterization of the gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1199-1203(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
RC   TISSUE=Larva;
RX   MEDLINE=93261414; PubMed=8492800;
RA   Poeck B., Balles J., Pflugfelder G.O.;
RT   "Transcript identification in the optomotor-blind locus of Drosophila
RT   melanogaster by intragenic recombination mapping and PCR-aided
RT   sequence analysis of lethal point mutations.";
RL   Mol. Gen. Genet. 238:325-332(1993).
CC   -!- FUNCTION: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
CC       REGULATOR. FLIES WITH L(1)OMB MUTATIONS SHOW SEVERE MALDEVELOPMENT
CC       OF THE OPTIC LOBES, REDUCTION IN WING SIZE AND AN INCREASED
CC       ABDOMINAL PIGMENTATION. THEY DIE DURING THE PUPAL STAGE.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: IN THIRD-INSTAR LARVAE IT IS FOUND IN THE
CC       BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
CC       THE THORACIC PART OF THE VENTRAL GANGLION.
CC   -!- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
CC       EMBRYOGENESIS, THE SECOND DAY OF PUPAL DEVELOPMENT AND IN THE
CC       ADULT.
CC   -!- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; M81796; AAA28736.1; -.
DR   EMBL; AE003431; AAF45946.1; -.
DR   EMBL; S61732; AAB26697.1; -.
DR   EMBL; S61727; AAB26697.1; JOINED.
DR   EMBL; S61729; AAB26697.1; JOINED.
DR   EMBL; S61744; AAB26699.1; -.
DR   EMBL; S61743; AAB26699.1; JOINED.
DR   EMBL; S61955; AAB26699.1; JOINED.
DR   HSSP; P24781; 1XBR.
DR   FlyBase; FBgn0000179; bi.
DR   InterPro; IPR001699; TF_T-box.
DR   Pfam; PF00907; T-box; 1.
DR   PRINTS; PR00937; TBOX.
DR   SMART; SM00425; TBOX; 1.
DR   PROSITE; PS01283; TBOX_1; 1.
DR   PROSITE; PS01264; TBOX_2; 1.
DR   PROSITE; PS50252; TBOX_3; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
FT   DOMAIN       51    140       ASN-RICH.
FT   DOMAIN      104    107       POLY-THR.
FT   DOMAIN      179    184       POLY-SER.
FT   DOMAIN      229    236       POLY-GLN.
FT   DOMAIN      238    244       POLY-PRO.
FT   DNA_BIND    332    513       T-BOX.
FT   DOMAIN      574    577       POLY-ASP.
FT   DOMAIN      607    692       ALA-RICH.
FT   DOMAIN      823    831       POLY-GLY.
FT   DOMAIN      910    916       POLY-ALA.
FT   DOMAIN      926    966       GLN/HIS-RICH.
FT   CONFLICT     10     10       F -> L (IN REF. 1 AND 3).
FT   CONFLICT    216    216       A -> P (IN REF. 1 AND 3).
FT   CONFLICT    511    511       F -> L (IN REF. 1).
FT   CONFLICT    823    823       MISSING (IN REF. 1).
FT   CONFLICT    976    988       MISSING (IN REF. 1).
SQ   SEQUENCE   988 AA;  103992 MW;  032B7A4471743FC9 CRC64;

  Query Match             4.9%;  Score 117.5;  DB 1;  Length 988;
  Best Local Similarity   35.2%;  Pred. No. 0.26;
  Matches   32;  Conservative   11;  Mismatches   23;  Indels   25;  Gaps    3;

Qy        372 NHGLVDSATINNCNTILENCRN--------------------SSDTTTTNNNNSVDRPSD 411
              | |  :| : || |:   |  |                    ||:||||:| |:    |
Db         59 NSGNNNSNSNNNTNSNTNNTNNLVAVSPTGGGAQLSPQSNHSSSNTTTTSNTNN----SS 114

Qy        412 SNTNNNNIVDHPNDINNKNNVDNKDNNSRDK 442
              || ||||   | |: |: || :| :||:  |
Db        115 SNNNNNNST-HNNNNNHTNNNNNNNNNTSQK 144


RESULT 10
HMCU_DROME
ID   HMCU_DROME     STANDARD;      PRT;  2175 AA.
AC   P10180; Q9W3Q6;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Homeobox protein cut.
GN   CT OR CG11387.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88232956; PubMed=2897632;
RA   Blochlinger K., Bodmer R., Jack J., Jan L.Y., Jan Y.N.;
RT   "Primary structure and expression of a product from cut, a locus
RT   involved in specifying sensory organ identity in Drosophila.";
RL   Nature 333:629-635(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: REGULATOR OF CELL FATE DECISIONS IN MULTIPLE LINEAGES.
CC       SPECIFICALLY, FUNCTIONS AS A DETERMINATION FACTOR THAT SPECIFIES
CC       SENSORY ORGAN IDENTITY IN PRECURSOR CELLS. PROBABLY ALSO INVOLVED
CC       IN CELL TYPE SPECIFICATION OF MALPIGHIAN TUBULES. IN ABSENCE OF
CC       CUT GENE EXTERNAL SENSORY ORGANS ARE TRANSFORMED INTO CHORDOTONAL
CC       ORGANS.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: DETECTED IN MANY CELLS IN THE CENTRAL NERVOUS
CC       SYSTEM, ALL EXTERNAL SENSORY ORGANS, SOME PERIPHERAL NEURONS, AND
CC       IN THE NON-NEURAL CELLS OF THE SPIRACLES AND THE MALPIGHIAN
CC       TUBULES.
CC   -!- DEVELOPMENTAL STAGE: CELL-SPECIFIC PATTERN OF EXPRESSION. BROADLY
CC       EXPRESSED DURING EMBRYONIC DEVELOPMENT.
CC   -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEOBOX MAY PARTICIPATE IN
CC       REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC       HETERODIMERIZATION.
CC   -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC   -!- SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
CC   ----------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------------
DR   EMBL; X07985; CAA30794.1; -.
DR   EMBL; AE003441; AAF46264.2; -.
DR   PIR; S03170; S03170.
```

```
DR   TRANSFAC; T02004; -.
DR   FlyBase; FBgn0004198; ct.
DR   InterPro; IPR000047; HTH_repressr.
DR   InterPro; IPR003350; Hmoeo_CUT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   Pfam; PF02376; CUT; 3.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Transcription regulation; Homeobox; DNA-binding;
KW   Developmental protein; Nuclear protein; Repeat; Coiled coil.
FT   DOMAIN      265    343       COILED COIL (POTENTIAL).
FT   DOMAIN      433    499       COILED COIL (POTENTIAL).
FT   DNA_BIND    798    964       CUT 1.
FT   DOMAIN     1056   1161       COILED COIL (POTENTIAL).
FT   DNA_BIND   1329   1417       CUT 2.
FT   DOMAIN     1463   1522       COILED COIL (POTENTIAL).
FT   DNA_BIND   1608   1695       CUT 3.
FT   DNA_BIND   1745   1804       HOMEOBOX.
FT   DOMAIN      194    210       ALA/GLN-RICH.
FT   DOMAIN      235    243       ALA-RICH.
FT   DOMAIN      271    293       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      384    428       ASN-RICH.
FT   DOMAIN      547    554       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      574    584       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      616    630       ALA-RICH.
FT   DOMAIN      665    699       HIS/GLN-RICH (OPA-REPEAT).
FT   DOMAIN     2004   2014       ALA-RICH.
FT   DOMAIN     2071   2077       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN     2124   2136       ALA/PRO-RICH.
SQ   SEQUENCE   2175 AA;  233628 MW;  08BF80C4861BD0AB CRC64;

  Query Match             4.9%;  Score 117;  DB 1;  Length 2175;
  Best Local Similarity   23.1%;  Pred. No. 0.8;
  Matches   58;  Conservative   32;  Mismatches  107;  Indels   54;  Gaps    8;

Qy        220 IPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMA 279
              ||     :::  ||     : |   ||   |  :  | |            |::
Db        222 IPTEQPAATVATGAAAAAAAAATPIATGNVKSGSTTSNAN----------HTNSNNSHQ 270

Qy        280 LEQVMSDRDSEDEVDDDVADFEDR-QMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHI 338
               |: : | : ::| |:|  | |:   |  :  |:  | :|      |   :|:     |
Db        271 DEEELDDEEEDEEEDEDEDDEEENASMQSNADDMELDAQQETRTEPSATTQQQ-----HQ 325

Qy        339 SWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSD-- 396
                |       |:  |              |   ||   ||   |:|:   |  | |:
Db        326 QQDTEDLEE--NKDAGEAS---------LNVSNNHNTTDSN--NSCSRKNNNGGNESEQH 372

Qy        397 ------------TTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNN----- 438
                          | |:|||:    : |||||||  :  :  : |    | :||
Db        373 VASSAEDDDCANNNTNTSNNNNTSNTATSNTNNNNNNNSSSGNSEKRKKKNNNNNNGQPA 432

Qy        439 ----SRDKVIK 445
                  ::|| ||
Db        433 VLLAAKDKEIK 443


RESULT 11
YIN0_YEAST
ID   YIN0_YEAST     STANDARD;      PRT;   964 AA.
AC   P40467;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Putative 108.8 kDa transcriptional regulatory protein in FKH1-STH1
DE   intergenic region.
GN   YIL130W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RA   Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA   Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA   Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA   Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA   Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA   Walsh S.V., Whitehead S.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC       CLUSTER DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z38059; CAA86148.1; -.
DR   PIR; S48404; S48404.
DR   HSSP; P08657; 1CLD.
DR   SGD; S0001392; YIL130W.
DR   InterPro; IPR001138; Fungi_TrN.
DR   Pfam; PF00172; Zn_clus; 1.
DR   PRINTS; PR00054; FUNGALZNCYS.
DR   SMART; SM00066; GAL4; 1.
DR   PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR   PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW   Hypothetical protein; Transcription regulation; DNA-binding;
KW   Nuclear protein; Zinc; Metal-binding.
FT   DNA_BIND     21     47       ZN(2)-CYS(6), FUNGAL-TYPE.
FT   DOMAIN      811    896       ASN-RICH.
SQ   SEQUENCE   964 AA;  108780 MW;  AD5ABE59E4B022CC CRC64;

  Query Match             4.8%;  Score 114;  DB 1;  Length 964;
  Best Local Similarity   40.4%;  Pred. No. 0.47;
  Matches   23;  Conservative   11;  Mismatches   21;  Indels    2;  Gaps    1;

Qy        382 NNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNN 438
              || | |  |  |: :    :|||: :  :::| ||||  :  |  || || :||:||
Db        833 NNNNNITNNSNNNHNNNNNDNNNNNNNNNNNNNNNNNNSGNSSN--NNNNNNNKNNN 887


RESULT 12
FAB1_YEAST
ID   FAB1_YEAST     STANDARD;      PRT;  2278 AA.
AC   P34756;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable phosphatidylinositol-4-phosphate 5-kinase FAB1 (EC 2.7.1.68)
DE   (1-phosphatidylinositol-4-phosphate kinase) (PIP5K) (PtdIns(4)P-5-
DE   kinase) (Diphosphoinositide kinase).
GN   FAB1 OR YFR019W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95392039; PubMed=7663021;
RA   Yamamoto A., Dewald D.B., Boronenkov I.V., Anderson R.A., Emr S.D.,
RA   Koshland D.;
RT   "Novel PI(4)P 5-kinase homologue, Fab1p, essential for normal vacuole
RT   function and morphology in yeast.";
RL   Mol. Biol. Cell 6:525-539(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=S288c / AB972;
RX   MEDLINE=95400292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
RN   [3]
RP   SIMILARITY TO TCP1/CCT FAMILY.
RX   MEDLINE=95314774; PubMed=7794526;
RA   Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RT   "Primary structure of the thermosome from Thermoplasma acidophilum.";
RL   Biol. Chem. Hoppe-Seyler 376:119-126(1995).
CC   -!- FUNCTION: CATALYSES THE PHOSPHORYLATION OF PHOSPHATIDYLINOSITOL-4-
CC       PHOSPHATE ON THE FIFTH HYDROXYL OF THE MYO-INOSITOL RING, TO FORM
CC       PHOSPHATIDYLINOSITOL-4,5-BIPHOSPHATE. REQUIRED FOR ENDOCYTIC-
CC       VACUOLAR PATHWAY AND NUCLEAR MIGRATION. THE PRODUCT OF THE
CC       REACTION IT CATALYSES FUNCTIONS AS AN IMPORTANT REGULATOR OF
CC       VACUOLE HOMEOSTASIS PERHAPS BY CONTROLLING MEMBRANE FLUX TO AND/OR
CC       FROM THE VACUOLE.
CC   -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
CC       monophosphate = ADP + 1-phosphatidyl-1D-myo-inositol 4,5-
CC       bisphosphate.
CC   -!- SUBCELLULAR LOCATION: VACUOLE-ASSOCIATED (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PTDINS(4)P-5-KINASE FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U01017; AAA81360.1; -.
DR   EMBL; D50617; BAA09258.1; -.
DR   SGD; S0001915; FAB1.
DR   InterPro; IPR002498; PIP5K.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF01363; FYVE; 1.
DR   Pfam; PF01504; PIP5K; 1.
DR   SMART; SM00064; FYVE; 1.
DR   SMART; SM00330; PIPKc; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
KW   Transferase; Kinase; Zinc-finger; Zinc.
FT   ZN_FING     240    299       FYVE-TYPE.
FT   DOMAIN      393    397       POLY-PRO.
FT   DOMAIN      571    590       POLY-ASN.
FT   DOMAIN     1808   1811       POLY-THR.
FT   DOMAIN     1891   1897       POLY-GLN.
FT   CONFLICT   2275   2275       R -> W (IN REF. 2).
SQ   SEQUENCE   2278 AA;  257417 MW;  1A0A30E13165DE41 CRC64;

  Query Match             4.7%;  Score 112.5;  DB 1;  Length 2278;
  Best Local Similarity   19.0%;  Pred. No. 1.9;
  Matches   88;  Conservative   57;  Mismatches  163;  Indels  155;  Gaps   18;

Qy         57 KRKRKSRSTGMVVFNY-------KDC------NNTLQKTEVREDCSCPFCSMLCGSFKGL 103
              ::| |: : |::   |       |:|       || ::      |   |||      |
Db        221 RKKDKTSNNGVLSKEYWMKDESSKECFSCGKTFNTFRRKHHCRICGQIFCSSCTLLIDGD 280

Qy        104 QFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR-RQ 162
              :|     |      :  :|   |             |: ||::      | ::|| | |
Db        281 RF---GCHAKMRVCYNCYEHADTY-----------EDSSDEENDSTMQL-NEPRSRSRS 324

Qy        163 RGGRNNTRRLKVCFLPLDSPSLTNGT-------------ENGITLLNDGNRGLGYPEAT 208
              |   |        | | |    |||              :| : |||| :      ::
Db        325 RSSNTNPYSHSHSHLHLISQDNHNGTDLHDPVAATDNPQQQNEVYLLNDDD-----VQSI 379

Qy        209 ELAGQ----FEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATK---TRKLSAERSEARSH 261
                :|:    |  |   || :|                 :|||   : ::| :    |:

Db        380 MTSGEDSKLFISTPPPPPKMA----------------IPATKQGGSLEISFDSENDRA- 421

Qy        262 LLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMH 321
              |  |       | :  :     : | |:    | :         |    : |      :
Db        422 LHYQDDNPGRHHHLDSVPTRYTIRDMDNISHYDTN-----SNSTLRPHYNTNNSTITINN 476

Qy        322 LWNS----------------------------FVRKQRVIADGHISWACEAFSRFYEKE 352
              | |:                           :|    |  | : | |  | |
Db        477 LNNTTSNNSNYNNTNSNSNINNPAHSLRRSIFHYVSSNSVNKDSNNSSATPASS------ 530

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENC--RN-------------SSD 396
                                    |: ::  | |: |   ||              :|
Db        531 ----------------------AQSSSILDPANRIIGNYAHRNYKFKFNYNSKGPSQQND 568

Qy        397 TTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNS 439
              |   ||:|: :  :::| ||||     : | : ||: : || :
Db        569 TANGNNDNNNNNNNNNNNNNNN---SASGIADNNNIPSNDNGT 608


RESULT 13
SWI1_YEAST
ID   SWI1_YEAST     STANDARD;      PRT;  1314 AA.
AC   P09547;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Transcription regulatory protein SWI1 (SWI/SNF complex component SWI1)
DE   (Transcription regulatory protein ADR6) (Regulatory protein GAM3).
GN   ADR6 OR SWI1 OR GAM3 OR YPL016W OR LPA1.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89057455; PubMed=3143101;
RA   O'Hara P.J., Horowitz H., Eichinger H., Young E.T.;
RT   "The yeast ADR6 gene encodes homopolymeric amino acid sequences and a
RT   potential metal-binding domain.";
RL   Nucleic Acids Res. 16:10153-10170(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313271; PubMed=9169875;
RA   Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA   Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA   Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA   Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA   Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA   Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA   Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA   Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA   Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA   Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA   Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA   Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA   Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA   Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA   Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA   Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL   Nature 387:103-105(1997).
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=92154671; PubMed=1339306;
RA   Peterson C.L., Herskowitz I.;
RT   "Characterization of the yeast SWI1, SWI2, and SWI3 genes, which
RT   encode a global activator of transcription.";
RL   Cell 68:573-583(1992).
CC   -!- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF
CC       COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER
CC       OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE
CC       BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.
```

```
CC   -!- SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR
CC       COMPLEX.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: CONTAINS 1 ARID DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See.http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X12493; CAA31013.1; -.
DR   EMBL; U33335; AAB68089.1; -.
DR   PIR; S05728; TNBYR6.
DR   TRANSFAC; T01279; -.
DR   SGD; S0005937; SWI1.
DR   InterPro; IPR001606; ARID.
DR   Pfam; PF01388; ARID; 1.
DR   SMART; SM00501; BRIGHT; 1.
KW   Transcription regulation; Activator; DNA-binding; Nuclear protein;
KW   Zinc-finger.
FT   DOMAIN        5     65       ASN/THR-RICH.
FT   DOMAIN      337    385       GLN-RICH.
FT   ZN_FING    1241   1258       C4-TYPE.
SQ   SEQUENCE   1314 AA;  147938 MW;  F442D5A82013CDBD CRC64;

  Query Match              4.7%;  Score 112;  DB 1;  Length 1314;
  Best Local Similarity   35.1%;  Pred. No. 1;
  Matches   33;  Conservative   11;  Mismatches   28;  Indels   22;  Gaps    5;

Qy        372 NHGLVDSATINNCNTILENCRNSSD-TTTTNNNNSVDRPS--DSNTNNNN-------IVD 421
              |:    : | || ||   |  |:::  ||||||   |  ::|||||       : |
Db         13 NNTTTTTTTTNNNNTNNNNTNNNNNPANNTNNNNSTGHSSNTNNNTNNNNTNTGASGVDD 72

Qy        422 HPN-----------DINNKN-NVDNKDNNSRDKV 443
              |           | || | | :| |||: : |
Db         73 FQNFFDPKPFDQNLDSNNNNSNSNNNDNNNSNTV 106


RESULT 14
CIGB_DICDI
ID   CIGB_DICDI     STANDARD;      PRT;   735 AA.
AC   Q94481;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Protein cigB (Fragment).
GN   CIGB.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Loomis W.F.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO D.DISCOIDEUM CIGA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66528; AAB06791.1; -.
DR   DictyDb; DD01058; cigB.
DR   InterPro; IPR000315; Znf_Bbox.
DR   Pfam; PF00643; zf-B_box; 1.
KW   Repeat.
```

```
FT   NON_TER       1      1
FT   DOMAIN      151    155       POLY-ASN.
FT   DOMAIN      251    268       POLY-ASN.
FT   DOMAIN      328    667       TANDEM REPEATS.
SQ   SEQUENCE   735 AA;  83443 MW;  9765013D88DA6C8C CRC64;

  Query Match              4.7%;  Score 111.5;  DB 1;  Length 735;
  Best Local Similarity   19.9%;  Pred. No. 0.52;
  Matches   73;  Conservative   56;  Mismatches  117;  Indels  121;  Gaps   17;

Qy         87 DCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDD 146
              ||    |  || | || : | : :  |                |    |:  :| | |
Db         20 DCKFMPCCSLCTSRKG-EHHGHKTDSL---------ESSASNIHSLINDF--------KD 61

Qy        147 KFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPE 206
              |    |  :|::   |:  ||           :| |   |  || |  |
Db         62 NIYP-----KLIERKE----NDQILLKE----------SNETFKEIQSQNDDNNNL---- 98

Qy        207 ATELAGQFEMTSNIPPAI-------AHSSLDAGAKV--ILTSEAVVPATKTRKLSAERSE 257
              |  :|:   ||  :        ::||     :  |:||          |: :
Db         99 ---LKNEFKQIHNILSFVELDIEKQLSTNLDQNILINTIITSS----INNDNKIISTILN 151

Qy        258 ARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEK 317
              :: :: | ::           |: : :|::|  |    | :  ::: :    :
Db        152 NNNNSIIDKLNYF----------TQINNQQDNDDNDDCYQIDTDTIELIKQY------QN 195

Qy        318 QFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLV- 376
              : | ||          :|:                                | :| :|: ::
Db        196 SLLILKNS----------NNIN----------------------NLTKLKEYNNQILK 221

Qy        377 -DSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNK 435
               |: ||||      ::  :  |      |:::|   ::| ||||              || || :|:
Db        222 FDNQTINNIKNSFKSIYSFGDIP----NDAIDYNINNNNNNNN--------NNNNNNNE 269

Qy        436 DNNSRDK 442
               ||| :|
Db        270 LNNSNNK 276


RESULT 15
CYAG_DICDI
ID   CYAG_DICDI     STANDARD;      PRT;   858 AA.
AC   Q03101;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Adenylate cyclase, germination specific (EC 4.6.1.1) (ATP
DE   pyrophosphate-lyase) (Adenylyl cyclase).
GN   ACGA OR ACG.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92233467; PubMed=1348970;
RA   Pitt G.S., Milona N., Borleis J., Lin K.C., Reed R.R.,
RA   Devreotes P.N.;
RT   "Structurally distinct and stage-specific adenylyl cyclase genes play
RT   different roles in Dictyostelium development.";
RL   Cell 69:305-315(1992).
RN   [2]
RP   3D-STRUCTURE MODELING OF 387-543.
RX   MEDLINE=98054247; PubMed=9391039;
RA   Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;
RT   "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT   and mutational analysis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC   -!- FUNCTION: ACG HAS A LARGE EXTRACELLULAR DOMAIN WHICH MAY BE
CC       INVOLVED IN THE RECOGNITION OF AN EXTRACELLULAR SIGNAL PRESENT
CC       DURING GERMINATION, LEADING TO ACTIVATION OR INHIBITION OF CAMP
CC       SYNTHESIS BY THE CYTOPLASMIC DOMAIN.
CC   -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC   -!- ENZYME REGULATION: INSENSITIVE TO GUANINE NUCLEOTIDES.
```

```
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- DEVELOPMENTAL STAGE: AFTER FRUITING BODIES HAVE BEEN FORMED AND
CC       DURING GERMINATION.
CC   -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC       FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 CHASE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M87278; AAA33164.1; -.
DR   PDB; 1AWM; 28-JAN-98.
DR   DictyDb; DD04007; acgA.
DR   InterPro; IPR001054; G_cyclase.
DR   Pfam; PF00211; guanylate_cyc; 1.
DR   SMART; SM00044; CYCc; 1.
DR   PROSITE; PS50839; CHASE; 1.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR   PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KW   Lyase; cAMP synthesis; Transmembrane; Germination; Signal-anchor;
KW   3D-structure.
FT   DOMAIN        1     18       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     19     41       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       42    858       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       86    317       CHASE.
FT   DOMAIN      396    526       GUANYLATE CYCLASE.
FT   DOMAIN      738    848       ASN-RICH.
FT   DOMAIN      835    848       POLY-ASN.
SQ   SEQUENCE   858 AA;  98413 MW;  B483FB255289E65D CRC64;

  Query Match             4.7%;  Score 111.5;  DB 1;  Length 858;
  Best Local Similarity   19.2%;  Pred. No. 0.64;
  Matches   76;  Conservative   57;  Mismatches  135;  Indels  127;  Gaps   15;

Qy        104 QFHLNSS-HDLFEFEFKLFEEYQTVNVSVK---LNSFIFEEEGSDDDKFEPFSLCSKPRK 159
              | |::   : | : :|   |    ::|  |      ::  :: ||    : ||     |
Db        533 QIHVSDRVYQLGKEDFNFSERCDIIHVKGKGRMKTWYLMGKKSSDFSLKKDFS-----RS 587

Qy        160 RRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSN 219
              | |                  ||| |  :: :  :        |||         : :|
Db        588 RVQ------------------PSLFNRKQSHVHCI--------YPEFPSGLQALNIENN 620

Qy        220 IPPAIAHSSLDAG----AKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRV 275
              :      :: |||    :|::  : |  |   |                   :||
Db        621 L------NNTDAGCENCSKILKKTYAYSPDHST----------------SNYYYHGDDN 657

Qy        276 QPMALEQVMSDR-DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMH--LWNSFVRKQRV 332
               |      :|  |  :  ||      :    |   |: :|| :     |:|    |::
Db        658 SPPPPSLNSNDLIDGSEYHDDPFPSDSNVGYHDTSKDIKEDENEQNETLLFNQEQLKKKQ 717

Qy        333 IADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR 392
              | :              :::|    |      :  ||: |:   ::   || |    |
Db        718 IEN-------------IQRDLSLNDS------IEAIKILNNNNNNNINDNNINNTNFNNN 758

Qy        393 ------NSSDTTTT--NNNNSVDRPSDSNTNNNNIVDH---------------------- 422
                   |:||      ||||  |:  ::: :|||| |:
Db        759 NNDININNSDNVNNYENNNNFSDKIENNDGDNNNINDNNYKSTNENNIKSKTLFKDSKSL 818

Qy        423 ------------PNDINNKNNVDNKDNNSRDKVIK 445
                         || || || :| :||: |: ::
Db        819 INDIKMAKENCDNNDDNNNNNNNNNNNNNNDENVE 853


Search completed: June 13, 2003, 15:36:11
Job time : 25 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:45:40 ; Search time 4487 Seconds
(without alignments)
11168.939 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1722	100.0	1722	6	AX032817	AX032817 Sequence
	2	1722	100.0	1722	6	AX032888	AX032888 Sequence
	3	1720.4	99.9	1722	6	AX032934	AX032934 Sequence
	4	1678.6	97.5	1700	8	AF284500	AF284500 Arabidops
	5	1646.6	95.6	1722	8	AF284501	AF284501 Arabidops
	6	1569.6	91.1	1715	6	AX032818	AX032818 Sequence
	7	1569.6	91.1	1715	6	AX032891	AX032891 Sequence
	8	1568.6	91.1	1721	8	AY034902	AY034902 Arabidops
	9	1537.6	89.3	1737	6	AX032894	AX032894 Sequence
	10	1277.4	74.2	1354	8	AY063047	AY063047 Arabidops
	11	412.2	23.9	6338	6	AX032890	AX032890 Sequence
c	12	407.8	23.7	91660	8	ATRPP5LE2	AF180942 Arabidops
	13	319.4	18.5	5895	6	AX032893	AX032893 Sequence
	14	319.4	18.5	199548	8	ATCHRIV45	AL161545 Arabidops
	15	319.4	18.5	201471	8	ATFCA7	Z97342 Arabidopsis
	16	291.4	16.9	2279	8	AB053171	AB053171 Arabidops
	17	236.2	13.7	2384	8	AF096096	AF096096 Arabidops
	18	230.8	13.4	558	6	AX032900	AX032900 Sequence
	19	163.8	9.5	325	6	AX032898	AX032898 Sequence
	20	126.6	7.4	463	6	AX032896	AX032896 Sequence
c	21	78.4	4.6	35276	2	AC115611	AC115611 Dictyoste
	22	75.2	4.4	256774	2	AC116964	AC116964 Dictyoste
c	23	73.4	4.3	106434	3	AC117080	AC117080 Dictyoste
c	24	73.4	4.3	147192	2	AC116925	AC116925 Dictyoste
c	25	73.2	4.3	79554	8	NCB11B23	AL669991 Neurospor
	26	72.8	4.2	141008	2	AC126897	AC126897 Rattus no
	27	72.8	4.2	144022	2	AC109949	AC109949 Rattus no
c	28	72.8	4.2	185424	2	AC096166	AC096166 Rattus no
	29	72.2	4.2	166918	2	AC128567	AC128567 Rattus no
	30	72.2	4.2	237563	2	AC102954	AC102954 Rattus no
	31	71.8	4.2	7921	3	AF153362	AF153362 Dictyoste
c	32	71.8	4.2	97683	2	AC116548	AC116548 Dictyoste
	33	71.8	4.2	212383	2	AC119030	AC119030 Rattus no
	34	71.6	4.2	57906	8	NCB7J19	AL389899 Neurospor
	35	71	4.1	656	3	AF453312	AF453312 Cotesia c
	36	71	4.1	8045	6	AX034425	AX034425 Sequence
c	37	71	4.1	156533	2	AC117070	AC117070 Dictyoste
	38	71	4.1	229502	2	AC126871	AC126871 Rattus no
c	39	70.6	4.1	153359	2	AC096792	AC096792 Rattus no
	40	70.6	4.1	165337	2	AC096212	AC096212 Rattus no
c	41	70.6	4.1	176351	2	AC117007	AC117007 Rattus no
	42	70.6	4.1	179380	2	AC120325	AC120325 Rattus no
	43	70.4	4.1	2790	3	DDI301670	AJ301670 Dictyoste
	44	70.4	4.1	6016	8	AY077730	AY077730 Podospora
	45	70.4	4.1	53440	2	AC117078	AC117078 Dictyoste

ALIGNMENTS

RESULT 1
AX032817
LOCUS AX032817 1722 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 57 from Patent WO0046358.
ACCESSION AX032817
VERSION AX032817.1 GI:10279793
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Dean,C., West,J. and Johanson,U.
TITLE Plant gene
JOURNAL Patent: WO 0046358-A 57 10-AUG-2000;
DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;

```
          JOHANSON URBAN (SE)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Landsberg erecta VRN2 cDNA"
BASE COUNT      538 a    339 c    348 g    497 t
ORIGIN

  Query Match           100.0%;  Score 1722;  DB 6;  Length 1722;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
              |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 2
AX032888
LOCUS       AX032888               1722 bp    DNA     linear   PAT 22-SEP-2000
DEFINITION  Sequence 1 from Patent WO0044918.
ACCESSION   AX032888
VERSION     AX032888.1  GI:10279825
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
```

```
  JOURNAL   Patent: WO 0044918-A 1 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      538 a    339 c    348 g    497 t
ORIGIN

  Query Match          100.0%;  Score 1722;  DB 6;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722
              |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722

RESULT 3
AX032934
LOCUS       AX032934               1722 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 47 from Patent WO0044918.
ACCESSION   AX032934
VERSION     AX032934.1  GI:10279842
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
```

```
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 47 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      539 a    339 c    347 g    497 t
ORIGIN

  Query Match           99.9%;  Score 1720.4;  DB 6;  Length 1722;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGAA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722
              |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 4
AF284500
LOCUS       AF284500               1700 bp    mRNA    linear   PLN 21-NOV-2001
DEFINITION  Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
            cds.
ACCESSION   AF284500
VERSION     AF284500.1  GI:16945787
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1700)
```

```
  AUTHORS   Gendall,A.R., Levy,Y.Y., Wilson,A. and Dean,C.
  TITLE     The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of
            Vernalization in Arabidopsis
  JOURNAL   Cell 107 (4), 525-535 (2001)
  MEDLINE   21575875
   PUBMED   11719192
REFERENCE   2  (bases 1 to 1700)
  AUTHORS   Gendall,A.R., Levy,Y.Y. and Dean,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre,
            Colney Lane, Norwich NR4 7UH, UK
FEATURES             Location/Qualifiers
     source          1. .1700
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="adjacent to RPP5"
                     /note="ecotype: Lansdberg erecta"
     gene            1. .1700
                     /gene="VRN2"
     CDS             227. .1564
                     /gene="VRN2"
                     /codon_start=1
                     /product="vernalization 2 protein"
                     /protein_id="AAL32135.1"
                     /db_xref="GI:16945788"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG
                     LQFHLNSSHDLFEFEFKLSEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
                     RQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTS
                     NIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
                     PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
                     DGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
                     NSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK"
     variation       853
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="a"
     variation       1195
                     /gene="VRN2"
                     /note="vrn2-1 mutation results in TGG Trp to TGA stop"
                     /replace="a"
     variation       1399
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="g"
     variation       1487
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="a"
     variation       1488
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="g"
BASE COUNT      533 a    334 c    347 g    486 t
ORIGIN

  Query Match             97.5%;  Score 1678.6;  DB 8;  Length 1700;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1681;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         20 TTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTTCAGGCCTCAATCCAAGACA 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 TTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTTCAGGCCTCAATCCAAGACA 75

Qy         80 TTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGCATTGAGTTTATCGCTAT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGCATTGAGTTTATCGCTAT 135

Qy        140 GACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTGCACTAACTTCATAATCGGC 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTGCACTAACTTCATAATCGGC 195

Qy        200 TCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGCAGAATTGTCGCGCGAAATC 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 TCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGCAGAATTGTCGCGCGAAATC 255

Qy        260 CTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCG 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCG 315

Qy        320 ACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTT 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 ACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTT 375

Qy        380 GAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTT 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTT 435

Qy        440 CAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTG 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTG 495

Qy        500 TCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCATC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCATC 555

Qy        560 TCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTC 619
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        556 TCATGATTTATTTGAATTTGAGTTCAAGCTTTCGGAAGAATACCAGACAGTTAATGTTTC 615

Qy        620 TGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCC 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCC 675

Qy        680 CTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 CTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAG 735

Qy        740 GAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAA 795

Qy        800 TGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCT 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 TGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCT 855

Qy        860 TGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGA 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        856 TGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGA 915

Qy        920 CGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAA 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 CGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAA 975

Qy        980 GTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTA 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 GTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTA 1035

Qy       1040 TCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGA 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 TCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGA 1095

Qy       1100 GGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGT 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 GGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGT 1155

Qy       1160 GGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACA 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 GGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACA 1215

Qy       1220 AAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGA 1279
```

```
Db      1216 AAGGGTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGA 1275

Qy      1280 GAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACT 1339

Db      1276 GAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACT 1335

Qy      1340 ATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAA 1399

Db      1336 ATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAA 1395

Qy      1400 TTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAG 1459

Db      1396 TTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAG 1455

Qy      1460 TGACTCAAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAA 1519

Db      1456 TGACTCAAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAA 1515

Qy      1520 CAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCG 1579

Db      1516 CAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCG 1575

Qy      1580 GCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTA 1639

Db      1576 GCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTA 1635

Qy      1640 GGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAA 1699

Db      1636 GGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAA 1695

Qy      1700 AAAAA 1704

Db      1696 AAAAA 1700
```

```
RESULT 5
AF284501
LOCUS       AF284501                1722 bp    mRNA    linear   PLN 21-NOV-2001
DEFINITION  Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
            cds, alternatively spliced.
ACCESSION   AF284501
VERSION     AF284501.1  GI:16945789
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Gendall,A.R., Levy,Y.Y., Wilson,A. and Dean,C.
  TITLE     The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of
            Vernalization in Arabidopsis
  JOURNAL   Cell 107 (4), 525-535 (2001)
  MEDLINE   21575875
   PUBMED   11719192
REFERENCE   2  (bases 1 to 1722)
  AUTHORS   Gendall,A.R., Levy,Y.Y. and Dean,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre,
            Colney Lane, Norwich NR4 7UH, UK
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="adjacent to RPP5"
                     /note="ecotype: Lansdberg erecta"
     gene            1. .1722
     CDS             227. .550
                     /gene="VRN2"
                     /gene="VRN2"
                     /note="alternatively spliced"
                     /codon_start=1
                     /product="vernalization 2 protein"
                     /protein_id="AAL32136.1"
                     /db_xref="GI:16945790"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKV
                     GNYYN"
BASE COUNT       540 a    338 c    353 g    491 t
ORIGIN
```

```
  Query Match             95.6%;  Score 1646.6;  DB 8;  Length 1722;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1681;  Conservative    0;  Mismatches    4;  Indels   22;  Gaps    1;

Qy        20 TTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTTCAGGCCTCAATCCAAGACA 79

Db        16 TTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTTCAGGCCTCAATCCAAGACA 75

Qy        80 TTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGCATTGAGTTTATCGCTAT 139

Db        76 TTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTTGCATTGAGTTTATCGCTAT 135

Qy       140 GACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTGCACTAACTTCATAATCGGC 199

Db       136 GACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTGCACTAACTTCATAATCGGC 195

Qy       200 TCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGCAGAATTGTCGCGCGAAATC 259

Db       196 TCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGCAGAATTGTCGCGCGAAATC 255

Qy       260 CTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCG 319

Db       256 CTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCG 315

Qy       320 ACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTT 379

Db       316 ACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTT 375

Qy       380 GAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTT 439

Db       376 GAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTT 435

Qy       440 CAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTG 499

Db       436 CAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTG 495

Qy       500 TCCATTTTGCTCTATGCTATGTGGTAGCTTCA----------------------AGGGGC 537

Db       496 TCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGGGC 555

Qy       538 TGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAG 597

Db       556 TGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTCGGAAG 615

Qy       598 AATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAA 657

Db       616 AATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAA 675

Qy       658 GTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAA 717

Db       676 GTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAA 735

Qy       718 GAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCA 777

Db       736 GAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCA 795

Qy       778 GTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAG 837

Db       796 GTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAG 855

Qy       838 GATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAG 897

Db       856 GATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAG 915
```

```
Qy        898 CCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGG 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        916 CCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGG 975

Qy        958 TCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTAC 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        976 TCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTAC 1035

Qy       1018 TTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1036 TTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAG 1095

Qy       1078 TAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1096 TAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATC 1155

Qy       1138 GCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1156 GCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 1215

Qy       1198 GGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTG 1257
              |||||||||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db       1216 GGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATTTCGTGGGCATGTG 1275

Qy       1258 AAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTT 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1276 AAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTT 1335

Qy       1318 GGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACA 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1336 GGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACA 1395

Qy       1378 ACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACA 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 ACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACA 1455

Qy       1438 ACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 ACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATC 1515

Qy       1498 CCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAG 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 CCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAG 1575

Qy       1558 TAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTC 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 TAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTC 1635

Qy       1618 TTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 TTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1695

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAA 1704
              ||||||||||||||||||||||||||
Db       1696 GATTACATTTTTAGTTAAAAAAAAAA 1722
```

```
RESULT 6
AX032818
LOCUS       AX032818               1715 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 58 from Patent WO0046358.
ACCESSION   AX032818
VERSION     AX032818.1  GI:10279794
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1715)
  AUTHORS   Dean,C., West,J. and Johanson,U.
  TITLE     Plant gene
  JOURNAL   Patent: WO 0046358-A 58 10-AUG-2000;
            DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
            JOHANSON URBAN (SE)
FEATURES             Location/Qualifiers
     source          1. .1715
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Columbia VRN2 cDNA"
BASE COUNT      536 a    326 c    353 g    500 t
ORIGIN
```

Query Match 91.1%; Score 1569.6; DB 6; Length 1715;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

```
Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||| ||||| |||||||||| |||| ||||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              |||||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718
              |||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720

Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838
              |||| |||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900
```

```
Qy        899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958
              ||||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy        959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy       1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy       1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy       1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy       1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy       1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318
              || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy       1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380

Qy       1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
              |||||||||||||||||||||||||||||||| |||||       ||| | |||||||||
Db       1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy       1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy       1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558
              ||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||
Db       1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy       1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618
              |||||||||||||   |||||| ||  |||||||||| ||||||||||||||||||||||
Db       1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy       1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
              ||||||||| |||||||||||||||||||||||||   |||||| ||||||||||||||||
Db       1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||   ||
Db       1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701
```

```
RESULT 7
AX032891
LOCUS       AX032891               1715 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 4 from Patent WO0044918.
ACCESSION   AX032891
VERSION     AX032891.1  GI:10279827
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1715)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 4 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1715
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      536 a    326 c    353 g    500 t
ORIGIN
```

Query Match 91.1%; Score 1569.6; DB 6; Length 1715;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

```
Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              |||||||||||||||||||||  |||||||||||||| ||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | | |||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||||| ||||| |||||||||||| |||| ||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              |||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718
              |||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720

Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838
              |||| |||||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900
```

```
Qy       899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958
Db       901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy       959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018
Db       961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy      1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078
Db      1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy      1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138
Db      1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy      1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198
Db      1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy      1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258
Db      1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy      1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318
Db      1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy      1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378
Db      1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380

Qy      1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
Db      1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy      1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498
Db      1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy      1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558
Db      1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy      1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618
Db      1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy      1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
Db      1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy      1678 GATTACATTTTTAGTTAAAAAAAAAAT 1705
Db      1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701
```

```
RESULT 8
AY034902
LOCUS       AY034902               1721 bp    mRNA    linear   PLN 23-APR-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
            cds.
ACCESSION   AY034902
VERSION     AY034902.1  GI:14334421
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1721)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1721)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G.,
            Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
            Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
            Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
            Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X.,
            Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D.,
            Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L.,
            Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E.,
            Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J.,
            Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1721
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL05-08-K19 (R09871)"
                     /note="This clone is in a modified pBluescript vector
                     (Lambda ZAP) as a XhoI/SstI insert.
                     ecotype: Columbia"
     gene            1. .1721
                     /gene="At4g16845"
     5'UTR           1. .231
                     /gene="At4g16845"
     CDS             232. .1554
                     /gene="At4g16845"
                     /codon_start=1
                     /evidence=experimental
                     /product="unknown protein"
                     /protein_id="AAK59409.1"
                     /db_xref="GI:14334422"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKG
                     LQFHLNSSHDLFEFEFKLLEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
                     RQRGGRNNTRRLKVCFLPLDSPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTS
                     NIPPAIAHSSLDAGAKVILTTEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
                     PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
                     DGHISWACEVFSRFYEKELHCYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
                     NTSVTNNNNNSVDHPSDSNTNNNNIVDHPNDIKNKNNVDNKDNNSRDK"
     3'UTR           1555. .1721
                     /gene="At4g16845"
BASE COUNT      548 a    324 c    353 g    496 t
ORIGIN

  Query Match           91.1%;  Score 1568.6;  DB 8;  Length 1721;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 1648;  Conservative    0;  Mismatches   49;  Indels   10;  Gaps    4;
```

```
Qy          2 AAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTT 59
              ||||||||||||||||||||  ||||||||||||| ||||||||||||||||||||||||
Db          1 AAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCTT 60

Qy         60 TCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGT 119
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGT 120

Qy        120 TGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTT 179
              ||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTT 180

Qy        180 GCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGG 239
              |||||||||||||||| ||||| |||||||||||| |||| ||||||||||||||||||
Db        181 GCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAGG 240

Qy        240 CAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTG 300

Qy        300 ATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCA 360

Qy        360 TCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGA 419
              ||||||||  ||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        361 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAGA 420

Qy        420 TCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAA 479
              |||||||||||||||||||||||||||||||||||||||||||  |||||||| | |||||||
Db        421 TCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAA 480

Qy        480 GTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTG 540

Qy        540 CAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAA 599
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        541 CAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGAA 600

Qy        600 TACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGT 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGT 660

Qy        660 GATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGA 719
              |||||  |||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 GATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAGA 720

Qy        720 GGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGT 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGT 780

Qy        780 TTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGA 839
              |||  ||||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGGA 840

Qy        840 TATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCC 899
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        841 TATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGCC 900

Qy        900 ATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTC 959
              |||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        901 ATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGTC 960

Qy        960 CCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTT 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTT 1020

Qy       1020 CTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTA 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTA 1080

Qy       1080 ATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGC 1139
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGC 1140

Qy       1140 CAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGG 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGG 1200

Qy       1200 AACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 1260

Qy       1260 GCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGG 1319
              | ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1261 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTGG 1320

Qy       1320 AGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAAC 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAAC 1380

Qy       1380 TGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAAC 1439
              |||||||||||||||||||||||||||||||| ||||       ||| | ||||||||||
Db       1381 TGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAAC 1434

Qy       1440 AACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCC 1499
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 AACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCG 1494

Qy       1500 AATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTA 1559
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||
Db       1495 AATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGTA 1553

Qy       1560 ATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTT 1619
              ||||||||||||  |||||| ||  |||||||||||| ||||||||||||||||||||||
Db       1554 ATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCTT 1613

Qy       1620 CTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTG 1678
              ||||||| |||||||||||||||||||||||  |||||| ||||||||||||||||||||
Db       1614 CTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACTG 1673

Qy       1679 ATTACATTTTTAGTTAAAAAAAAAAAT 1705
              ||||||||||||||||||||||   ||
Db       1674 ATTACATTTTTAGTTAAAAAAATGGAT 1700
```

```
RESULT 9
AX032894
LOCUS       AX032894               1737 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 7 from Patent WO0044918.
ACCESSION   AX032894
VERSION     AX032894.1  GI:10279829
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1737)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 7 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1737
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      543 a    330 c    359 g    505 t
ORIGIN

  Query Match           89.3%;  Score 1537.6;  DB 6;  Length 1737;
  Best Local Similarity   95.3%;  Pred. No. 0;
```

```
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   32;  Gaps    5;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              ||||||||||||||||||||||  |||||||||||| ||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||||| ||||| |||||||||||| |||| ||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              |||||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCA------- 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGTGGG 540

Qy        532 ---------------AGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 576
                             ||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 600

Qy        577 TTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 636
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 660

Qy        637 TCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 696
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        661 TCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 720

Qy        697 AACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 756
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 780

Qy        757 TTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTA 816
              |||||||||||||||||||||||||| ||||||||||||||||||||  |||| || | |
Db        781 TTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGA 840

Qy        817 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 900

Qy        877 TGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 936
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 960

Qy        937 TATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 996
              |||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 1020

Qy        997 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1080

Qy       1057 AGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACG 1116
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1081 AGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACG 1140

Qy       1117 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1200

Qy       1177 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1260

Qy       1237 GTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTT 1296
              |||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db       1261 GTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTT 1320

Qy       1297 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1380

Qy       1357 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAG 1416
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1381 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG 1440

Qy       1417 ACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACA 1476
                  ||| | |||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1441 ------TCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACA 1494

Qy       1477 ACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGG 1536
              |||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||||
Db       1495 ACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGG 1554

Qy       1537 ACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGAT 1596
              |||||||||||||||||| ||||||||||||||||||  |||||| ||  ||||||||||
Db       1555 ACAATAACAGCAGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGAT 1613

Qy       1597 TTATCGGATTGTAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTT 1655
                |||||||||||||||||||||||||||| |||||||||||||||||||||||||  ||
Db       1614 CTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATT 1673

Qy       1656 TTATATGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||| ||||||||||||||||||||||||||||||||||||||||   ||
Db       1674 TTATTTGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 1723
```

```
RESULT 10
AY063047
LOCUS       AY063047               1354 bp    mRNA    linear   PLN 18-APR-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
            cds.
ACCESSION   AY063047
VERSION     AY063047.1  GI:17104664
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1354)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Etgu,P.,
            Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H.,
            Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
            Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
            Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1354)
  AUTHORS   Yamada,K., Banh,J., Banno,F., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H.,
            Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J.,
```

```
            Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Banno,F., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T.,
            Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C.,
            Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A.,
            Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
            to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
            /PGEC) contributed equally to this work as PIs.
FEATURES              Location/Qualifiers
     source           1. .1354
                      /organism="Arabidopsis thaliana"
                      /db_xref="taxon:3702"
                      /chromosome="4"
                      /clone="U09871"
                      /note="This clone is in pUNI 51.
                      ecotype: Columbia"
     gene             1. .1354
                      /gene="At4g16845"
     CDS              1. .1323
                      /gene="At4g16845"
                      /codon_start=1
                      /evidence=experimental
                      /product="unknown protein"
                      /protein_id="AAL34221.1"
                      /db_xref="GI:17104665"
                      /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                      FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKG
                      LQFHLNSSHDLFEFEFKLLEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
                      RQRGGRNNTRRLKVCFLPLDSPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTS
                      NIPPAIAHSSLDAGAKVILTTEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
                      PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
                      DGHISWACEVFSRFYEKELHCYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
                      NTSVTNNNNNSVDHPSDSNTNNNNIVDHPNDIKNKNNVDNKDNNSRDK"
     3'UTR            1324. .1354
                      /gene="At4g16845"
BASE COUNT      431 a    264 c    287 g    372 t
ORIGIN

  Query Match           74.2%;  Score 1277.4;  DB 8;  Length 1354;
  Best Local Similarity   97.2%;  Pred. No. 2.4e-272;
  Matches 1323;  Conservative    0;  Mismatches   31;  Indels    7;  Gaps    2;

Qy        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 60

Qy        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 120

Qy        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410
              |||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
Db        121 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 180

Qy        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        181 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 240

Qy        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 300

Qy        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 360

Qy        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 420

Qy        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 480

Qy        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 540

Qy        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830
              |||||||||||| |||||||||||||||||||||  |||| || ||||||||||||||||
Db        541 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 600

Qy        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        601 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 660

Qy        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||| ||||
Db        661 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 720

Qy        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 780

Qy       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 840

Qy       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 900

Qy       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 960

Qy       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1020

Qy       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310
              |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1021 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1080

Qy       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1140

Qy       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430
              |||||||||||||||||||||||||||||||||||| |||||       ||| | | |
Db       1141 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAAC 1194

Qy       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1195 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1254

Qy       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550
              |||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1255 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1314
```

```
Qy       1551 GACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATA 1591
              ||| ||||||||||||||||||  |||||| ||  ||||||
Db       1315 GAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATA 1354


RESULT 11
AX032890
LOCUS       AX032890               6338 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 3 from Patent WO0044918.
ACCESSION   AX032890
VERSION     AX032890.1  GI:10279826
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 6338)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 3 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .6338
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT     1932 a   1076 c   1134 g   2193 t      3 others
ORIGIN

  Query Match           23.9%;  Score 412.2;  DB 6;  Length 6338;
  Best Local Similarity   98.8%;  Pred. No. 8.7e-81;
  Matches  414;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1304 ACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363
              | | | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5102 AATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 5161

Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5162 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 5221

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5222 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 5281

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543
              ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5282 CATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 5341

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5342 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 5401

Qy       1604 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5402 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 5461

Qy       1664 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5462 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 5520


RESULT 12
ATRPP5LE2/c
LOCUS       ATRPP5LE2              91660 bp    DNA     linear   PLN 22-DEC-2000
DEFINITION  Arabidopsis thaliana RPP5 disease resistance locus, main contig.
ACCESSION   AF180942 U97106
VERSION     AF180942.1  GI:6449044
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 76020 to 83095)
  AUTHORS   Parker,J.E., Coleman,M.J., Szabo,V., Frost,L.N., Schmidt,R., van
            der Biezen,E.A., Moores,T., Dean,C., Daniels,M.J. and Jones,J.D.
  TITLE     The Arabidopsis downy mildew resistance gene RPP5 shares similarity
            to the toll and interleukin-1 receptors with N and L6
  JOURNAL   Plant Cell 9 (6), 879-894 (1997)
  MEDLINE   97355983
   PUBMED   9212464
REFERENCE   2  (bases 1 to 91660)
  AUTHORS   Noel,L., Moores,T.L., van Der Biezen,E.A., Parniske,M.,
            Daniels,M.J., Parker,J.E. and Jones,J.D.
  TITLE     Pronounced intraspecific haplotype divergence at the RPP5 complex
            disease resistance locus of Arabidopsis
  JOURNAL   Plant Cell 11 (11), 2099-2112 (1999)
  MEDLINE   20025650
   PUBMED   10559437
REFERENCE   3  (bases 76020 to 83095)
  AUTHORS   Coleman,M.J., Parker,J.E., Szabo,V., van der Biezen,E.A.,
            Daniels,M.J. and Jones,J.D.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-1997) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich, Norfolk NR4 7UH, UK
REFERENCE   4  (bases 1 to 91660)
  AUTHORS   Noel,L., Moores,T., van der Biezen,E.A., Parniske,M., Daniels,M.J.,
            Parker,J.E. and Jones,J.D.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-1999) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich, Norfolk NR4 7UH, United Kingdom
COMMENT     On Nov 18, 1999 this sequence version replaced gi:2109274.
FEATURES             Location/Qualifiers
     source          1. .91660
                     /organism="Arabidopsis thaliana"
                     /cultivar="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /map="C18a"
     gene            5354. .9673
                     /gene="La-I"
                     /pseudo
     CDS             join(5354. .5829,5977. .7057,7241. .7510,7610. .8071,
                     8281. .9366,9458. .9673)
                     /gene="La-I"
                     /note="RPP5 homolog"
                     /pseudo
                     /codon_start=1
     repeat_region   15312. .20023
                     /rpt_family="retroelement castor"
                     /rpt_type=dispersed
     LTR             15312. .15431
     gene            15507. .19886
                     /gene="encodes gag/pol"
                     /pseudo
     LTR             19904. .20023
     gene            20048. .22736
                     /gene="La-H"
                     /pseudo
                     /evidence=not_experimental
     CDS             join(20048. .20440,20985. .22416,22533. .22736)
                     /gene="La-H"
                     /note="RPP5 homolog"
                     /pseudo
                     /codon_start=1
     gene            25823. .35584
                     /gene="La-G"
                     /note="interrupted by retroelement pollux1"
                     /pseudo
                     /evidence=not_experimental
     CDS             join(25823. .26301,26428. .27343,32765. .32926,33111. .33407,
                     33525. .33899,34108. .35235,35366. .35584)
```

```
                     /gene="La-G"
                     /note="RPP5 homolog"
                     /pseudo
     repeat_region   /codon_start=1
                     27344. .32762
                     /rpt_family="retroelement pollux1"
     LTR             /rpt_type=dispersed
     gene            27344. .27791
                     27937. .32166
                     /gene="encodes gag/pol"
     LTR             /pseudo
     gene            32315. .32762
                     36424. .39175
                     /gene="La-F"
                     /pseudo
     CDS             /evidence=not_experimental
                     join(36424. .37022,37167. .38250,38389. .38695,38855. .39175)
                     /gene="La-F"
                     /note="RPP5 homolog"
                     /pseudo
     gene            /codon_start=1
                     39222. .43288
                     /gene="La-E"
                     /pseudo
     CDS             /evidence=not_experimental
                     join(39222. .39754,39899. .40064,42749. .43288)
                     /gene="La-E"
                     /note="RPP5 homolog"
                     /pseudo
     gene            /codon_start=1
                     46918. .56304
                     /gene="La-D"
                     /note="interrupted by retroelement pollux2"
                     /pseudo
     CDS             /evidence=not_experimental
                     join(46918. .47393,47558. .48635,48820. .48990,54412. .54573,
                     54758. .55054,55172. .55546,55648. .55950,56080. .56304)
                     /gene="La-D"
                     /note="RPP5 homolog"
                     /pseudo
     repeat_region   /codon_start=1
                     48991. .54409
                     /rpt_family="retroelement pollux2"
     LTR             /rpt_type=dispersed
     gene            48991. .49438
                     49584. .53813
     LTR             /gene="encodes gag/pol"
     misc_feature    53962. .54409
                     complement(58189. .58275)
     misc_feature    /note="similar to 5S ribosomal RNA gene repeat unit"
                     complement(58305. .58773)
     gene            /note="similar to 5S ribosomal RNA gene repeat unit"
                     59720. .65992
                     /gene="La-C"
                     /pseudo
     CDS             /evidence=not_experimental
                     join(59720. .60183,60265. .61340,62850. .63143,63293. .63703,
                     63820. .64299,64540. .65646,65777. .65992)
                     /gene="La-C"
                     /note="RPP5 homolog"
                     /pseudo
     gene            /codon_start=1
                     68136. .72604
                     /gene="La-B"
                     /pseudo
     CDS             /evidence=not_experimental
                     join(68136. .68603,68739. .69793,69993. .70280,70398. .70870,
                     71092. .72240,72368. .72604)
                     /gene="La-B"
                     /note="RPP5 homolog"
                     /pseudo
     gene            /codon_start=1
                     <77312. .>82401

     mRNA            /gene="RPP5"
                     /evidence=experimental
                     join(<77312. .77790,77917. .78991,79181. .79474,79581. .79994,
                     80101. .80598,80811. .81842,81970. .82180,82319. .>82401)
                     /gene="RPP5"
     CDS             /product="RPP5"
                     join(77312. .77790,77917. .78991,79181. .79474,79581. .79994,
                     80101. .80598,80811. .81842,81970. .82180,82319. .82401)
                     /gene="RPP5"
                     /codon_start=1
                     /product="downy mildew resistance protein RPP5"
                     /protein_id="AAF08790.1"
                     /db_xref="GI:6449046"
                     /translation="MAASSSSGRRRYDVFPSFSGVDVRKTFLSHLLKALDGKSINTFI
                     DHGIERSRTIAPELISAIREARISIVIFSKNYASSTWCLNELVEIHKCFNDLGQMVIP
                     VFYDVDPSEVRKQTGEFGKVFEKTCEVSKDKQPGDQKQRWVQALTDIANIAGEDLLNG
                     PNEAHMVEKISNDVSNKLITRSKCFDDFVGIEAHIEAIKSVLCLESKEARMVGIWGQS
                     GIGKSTIGRALFSQLSSQFHHRAFLTYKSTSGSDVSGMKLSWQKELLSEILGQKDIKI
                     EHFGVVEQRLNHKKVLILLDDVDNLEFLKTLVGKAEWFGSGSRIIVITQDRQLLKAHE
                     IDLVYEVKLPSQGLALKMISQYAFGKDSPPDDFKELAFEVAELVGSLPLGLSVLGSSL
                     KGRDKDEWVKMMPRLRNDSDDKIEETLRVGYDRLNKKNRELFKCIACFFNGFKVSNVK
                     ELLEDDVGLTMLAEESLIRITPVGYIEMHNLLEKLGREIDRAKSKGNPGKRQFLTNFE
                     DIREVLTEKTGTETLLGIRLPHPGYLTTRSFLIDEKSFKGMRNLQYLEIGYWSDGVLP
                     QSLVYFPRKLKRLWWDNCPLKRLPSNFKAEYLVELRMVNSKLEKLWDGTQPLGSLKKM
                     DLYNSYKLKEIPDLSLAINLEELNLEECESLETLPSSIQNAIKLRELNCWGGLLIDLK
                     SLEGMCNLEYLSVPSWSSRECTQGIVYFPRKLKSVLWTNCPLKRLPSNFKAEYLVELI
                     MEYSELEKLWDGTQSLGSLKEMNLRYSNNLKEIPDLSLAINLEELDLFGCVSLVTLPS
                     SIQNATKLIYLDMSECENLESFPTVFNLKSLEYLDLTGCPNLRNFPAIKMGCAWTRLS
                     RTRLFPEGRNEIVVEDCFWNKNLPAGLDYLDCLMRCMPCEFRSEQLTFLNVSGCKLEK
                     LWEGIQSLGSLEEMDLSESENLKELPDLSKATNLKLLCLSGCKSLVTLPSTIGNLQNL
                     RRLYMNRCTGLEVLPTDVNLSSLETLDLSGCSSLRTFPLISTNIVCLYLENTAIEEIP
                     DLSKATKLESLILNNCKSLVTLPSTIGNLQNLRRLYMNRCTGLELLPTDVNLSSLETL
                     DLSGCSSLRTFPLISTRIECLYLENTAIEEVPCCIEDFTRLTVLRMYCCQRLKNISPN
                     IFRLTSLTLADFTDCRGVIKALSDATVVATMEDHVSCVPLSENIEYTCERFWDACSDY
                     YSDDFEVNRNPIRLSTMTVNDVEFKFCCSITIKECGVRLLYVYQETEHNQQTTRSKKR
                     MRMTSGTSEEDINLPYGQIVADTGLAALNTELSLGQGEASSSTSLEGEALCVDDYMIN
                     EEQDEQIPILYPVYDIDDDMWRSLFFGDTDDDMWRSLYSAE"
BASE COUNT     28475 a   17266 c   17533 g   28386 t
ORIGIN
```

```
  Query Match             23.7%;  Score 407.8;  DB 8;  Length 91660;
  Best Local Similarity   98.3%;  Pred. No. 8.2e-80;
  Matches  412;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1304 ACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363
             | | | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85502 AATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363

Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85442 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGGCGTAATAGCTCAGACACCAC 1423

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85382 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543
             |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85322 CATTGTGAGTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85262 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603

Qy       1604 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85202 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663

Qy       1664 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85142 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 85084
```

```
RESULT 13
AX032893
LOCUS       AX032893               5895 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 6 from Patent WO0044918.
ACCESSION   AX032893
VERSION     AX032893.1  GI:10279828
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 5895)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 6 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .5895
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT     1742 a   1027 c   1070 g   2056 t
ORIGIN

  Query Match             18.5%;  Score 319.4;  DB 6;  Length 5895;
  Best Local Similarity   92.7%;  Pred. No. 3.1e-60;
  Matches  370;  Conservative    0;  Mismatches   21;  Indels    8;  Gaps    3;

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4677 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 4736

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||| |||||       |||
Db       4737 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 4790

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       4791 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 4850

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db       4851 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 4910

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||| |||||||||||||||||  |||||| ||  |||||||||| ||||||||||
Db       4911 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 4969

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAG 1666
              |||||||||||||||||| |||||||||||||||||||||||  |||||| ||||||
Db       4970 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 5029

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||||   ||
Db       5030 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 5068


RESULT 14
ATCHRIV45
LOCUS       ATCHRIV45              199548 bp    DNA     linear   PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 45.
ACCESSION   AL161545
VERSION     AL161545.2  GI:7268431
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 199548)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV44 at the 5' end and an
            overlap with ATCHRIV46 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .199548
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     gene            6008. .10944
                     /gene="AT4g16810"
     gene            complement(join(6008. .6649,6855. .7090,7222. .7258,
                     7337. .7536,7629. .7772,7877. .7913,8744. .8853,9020. .9108,
                     9487. .9576,9704. .9786,9891. .9980,10178. .10229,
                     10883. .10944))
                     /gene="AT4g16810"
     CDS             complement(join(6008. .6649,6855. .7090,7222. .7258,
                     7337. .7536,7629. .7772,7877. .7913,8744. .8853,9020. .9108,
                     9487. .9576,9704. .9786,9891. .9980,10178. .10229,
                     10883. .10944))
                     /gene="AT4g16810"
                     /note="similarity to centromere protein E, human"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB80952.1"
                     /db_xref="GI:7268432"
                     /translation="MDPIKLTTEAKVPAKRSKATSHYLPLHKRQFYHSRTGQPLSLEQ
                     VMSDRDSENDVDKNDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAH
                     IPWACEAFSRLHLQELRSNLSLDLGNTETTETAPVATEATTHGEQDGIDVGEQEAVVV
                     EVVVAEAEVVVVVKAEAQAVVETSSSDLRWAGEDGYSDSTSPISEPPDITNWFPSYVY
                     ESLPLSSSINEIEIEVDVVNKANPGGVIDESLPSEPPDLGKWFSSYVYESPLLDTSDG
                     LELSVPGESECVKETQTENESPKIEGNDVCPRLLEQELVSSTKVTDFSESQSVLSEPP
                     DLRNWFSSYEYQSPQLSDIQEFGFLYSEKDELIIEESDTEEGISSGIFRKIKSRQETV
                     GLGRLDSNDYKENIATDTAKEVSLDNAVSDQKMEKRSSVRLLNASKKDVKQESSFKQE
                     PLLCELQEEARFSPRVSRYNPKLKSPSKNDTSLHELRPIHIQESISMNSNRQKSPPIE
                     QESYDKENVHGQSNQTGFVTMKKASFREARDQCSLKKPNTEVLVKCSSSKELKTKAGE
                     DNTEEREKKRRVLGEMSNQLSSVAKEIAGKWRCPQKNKRKIVPPLKQLRLDAWIHKV"
     exon            complement(6008. .6649)
                     /gene="AT4g16810"
                     /number=1
     intron          complement(6650. .6854)
                     /gene="AT4g16810"
                     /number=1
     exon            complement(6855. .7090)
                     /gene="AT4g16810"
                     /number=2
     intron          complement(7091. .7221)
                     /gene="AT4g16810"
                     /number=2
     exon            complement(7222. .7258)
                     /gene="AT4g16810"
                     /number=3
     intron          complement(7259. .7336)
                     /gene="AT4g16810"
                     /number=3
     exon            complement(7337. .7536)
                     /gene="AT4g16810"
                     /number=4
     intron          complement(7537. .7628)
                     /gene="AT4g16810"
                     /number=4
     exon            complement(7629. .7772)
                     /gene="AT4g16810"
                     /number=5
```

```
     intron          complement(7773. .7876)
                     /gene="AT4g16810"
                     /number=5
     exon            complement(7877. .7913)
                     /gene="AT4g16810"
                     /number=6
     intron          complement(7914. .8743)
                     /gene="AT4g16810"
                     /number=6
     exon            complement(8744. .8853)
                     /gene="AT4g16810"
                     /number=7
     intron          complement(8854. .9019)
                     /gene="AT4g16810"
                     /number=7
     exon            complement(9020. .9108)
                     /gene="AT4g16810"
                     /number=8
     intron          complement(9109. .9486)
                     /gene="AT4g16810"
                     /number=8
     exon            complement(9487. .9576)
                     /gene="AT4g16810"
                     /number=9
     intron          complement(9577. .9703)
                     /gene="AT4g16810"
                     /number=9
     exon            complement(9704. .9786)
                     /gene="AT4g16810"
                     /number=10
     intron          complement(9787. .9890)
                     /gene="AT4g16810"
                     /number=10
     exon            complement(9891. .9980)
                     /gene="AT4g16810"
                     /number=11
     intron          complement(9981. .10177)
                     /gene="AT4g16810"
                     /number=11
     exon            complement(10178. .10229)
                     /gene="AT4g16810"
                     /number=12
     intron          complement(10230. .10882)
                     /gene="AT4g16810"
                     /number=12
     exon            complement(10883. .10944)
                     /gene="AT4g16810"
                     /number=13
     gene            16254. .19247
                     /gene="AT4g16820"
     CDS             join(16254. .17790,18979. .19247)
                     /gene="AT4g16820"
                     /note="similarity to triacylglycerol lipase (EC 3.1.1.3),
                     Rhizopus niveus
                     Contains Lipases, serine active site AA321-330
                     contains EST gb:AI993282.1, T76589, AI994595.1, R30062,
                     Z32589"
                     /codon_start=1
                     /product="triacylglycerol lipase like protein"
                     /protein_id="CAB80953.1"
                     /db_xref="GI:7268433"
                     /translation="MQTLTPNADIFHAKRRRFTCNTHTSTLIPTKPLSVSPARKTNKE
                     HLRNLENVLRTSSNSIDHIENVTSRQEKTTKNTSTSSLLGGLNLARIWPQMKAAVDEM
                     SPKNLKRLQRLLSKSSEERSPKSKLGSKWRELHGLNNWAGLLDPLDENLRRELVRYGE
                     FVQAAYHAFHSDPEGSPRHVALPDGSFKVTKSLYATSSVRLPKWIDDVAPDLRWMTKQ
                     TSWVGYVAVCDDPREIRRMGRREIVIALRGTATLLEWSENFRPNLVSMPEPKPDQSDP
                     TRPKVECGFNSLYTTGDQHAPSLAESLVGEISRLVELYAGEELSISVTGHSLGAAIAL
                     LAADDIAERVPHAPPVAVFSFGGPRVGNREFADRLDSKGVKVLRVVNSQDVVTKVPGI
                     FADNDKQGQSRNNGRSPGGIMEMVERNNPWAYSHVGAELRVDMKMSPYLKPNADVACC
                     HDLEAYLHLVDGFLASNCPFRANAKRSLRKLLDEQRSNVKVLYTGKSLRLNRTFHDNG
                     DVLPRDTGQGDNPIMKYKLLFRYAVGLSILIITTVFLPIPNFHTFYFHFLIFISLSFL
     exon            SLTNQQVLLSTHFVSFILTFSLDLLIIPSMAWSSH"
                     16254. .17790
                     /gene="AT4g16820"
                     /number=1
     intron          17791. .18978
                     /gene="AT4g16820"
                     /number=1
     exon            18979. .19247
                     /gene="AT4g16820"
                     /number=2
     gene            19312. .20999
                     /gene="AT4g16830"
     CDS             join(19312. .19483,19652. .19958,20041. .20106,20210. .20336,
                     20437. .20592,20679. .20729,20811. .20999)
                     /gene="AT4g16830"
                     /note="cognate cDNA CDNA4KMY is matching at position
                     1428395-1430523
                     similarity to nuclear antigen - Bos taurus
                     contains EST gb:AI994595.1, AI993282.1, R30062, T42546,
                     Z32590, Z32589, T76589, AA395095"
                     /codon_start=1
                     /product="nuclear antigen homolog"
                     /protein_id="CAB80954.1"
                     /db_xref="GI:7268434"
                     /translation="MATLNPFDLLDDDAEDPSQLAVAIEKIDKSKKSGQVSSLPAKSA
                     PKLPSKPLPPAQAVREARSDAPRGGGGRGGFNRGRGGYNRDDGNNGYSGGYTKPSGEG
                     DVSKSSYERRGGGGAPRGSFRGEGGGPGGGRRGGFSNEGGDGERPRRAFERRSGTGRG
                     SDFKRDGSGRGNWGTPGEEIAAETEAVAGVETEKDVGEKPAVDDVAADANKEDTVVEE
                     KEPEDKEMTLDEYEKILEEKKKALQSLTTSERKVDTKVFESMQQLSNKKSNDEIFIKL
                     GSDKDKRKDDKEEKAKKAVSINEFLKPAEGGNYYRGGRGGRGRGGRGRGGVSSGESGG
     exon            YRNEAAPAIGDAAQFPSLGGK"
                     19312. .19483
                     /gene="AT4g16830"
                     /number=1
     intron          19484. .19651
                     /gene="AT4g16830"
                     /number=1
     exon            19652. .19958
                     /gene="AT4g16830"
                     /number=2
     intron          19959. .20040
                     /gene="AT4g16830"
                     /number=2
     exon            20041. .20106
```

```
Query Match          18.5%;  Score 319.4;  DB 8;  Length 199548;
Best Local Similarity  92.7%;  Pred. No. 3.1e-60;
Matches  370;  Conservative    0;  Mismatches   21;  Indels    8;  Gaps    3;

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28171 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 28230

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||||||||| |||||     |||
Db      28231 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 28284

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      28285 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 28344

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db      28345 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 28404

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||.|||||||||||||||  |||||| ||  |||||||||| ||||||||||||||
Db      28405 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 28463

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAG 1666
              |||||||||||||||||| |||||||||||||||||||||||  |||||| ||||||
Db      28464 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 28523

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||  ||
```

```
Db      28524 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 28562


RESULT 15
ATFCA7
LOCUS       ATFCA7                201471 bp    DNA     linear   PLN 29-JUN-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment
            No. 7.
ACCESSION   Z97342
VERSION     Z97342.2  GI:5302800
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 201471)
  AUTHORS   Bevan,M., Stiekema,W., Murphy,G., Wambutt,R., Pohl,T., Terryn,N.,
            Kreis,M., Kavanagh,T., Entian,K.D., Rieger,M., James,R.,
            Puigdomenech,P., Hatzopoulos,P., Obermaier,B., Duesterhoft,A.,
            Jones,J., Palme,K., Ansorge,W., Delseny,M., Bancroft,I.,
            Mewes,H.W., Schueller,C. and Chalwatzis,N.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201471)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     On Jun 30, 1999 this sequence version replaced gi:2245031.
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATFCA6 at the 5' end and an
            overlap with ATFCA8 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .201471
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     source          1. .8956
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB5K23"
     source          21108. .37844
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB6C5"
     source          37120. .39085
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB4A23"
     source          37904. .128469
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="BAC IGF5D3"
     source          38368. .63818
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC33M3"
     source          106243. .128469
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC16N19"
     source          113141. .132730
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC20I2"
     source          129847. .157017
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC1B24"
     source          152677. .190349
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid G16599"
     source          185303. .201471
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid G11319"
     gene            12153. .13466
                     /gene="dl4415w"
     CDS             join(12153. .12312,12416. .12714,12983. .13062,13151. .13466)
                     /gene="dl4415w"
                     /note="contains EST gb:
                     1216777;935272;906153;947895;315626"
                     /codon_start=1
                     /product="DNA-binding homeotic protein Athb-2"
                     /protein_id="CAB10452.1"
                     /db_xref="GI:2245033"
                     /db_xref="SWISS-PROT:Q05466"
                     /translation="MMFEKDDLGLSLGLNFPKKQINLKSNPSVSVTPSSSSFGLFRRS
                     SWNESFTSSVPNSDSSQKETRTFIRGIDVNRPPSTAEYGDEDAGVSSPNSTVSSSTGK
                     RSEREEDTDPQGSRGISDDEDGDNSRKKLRLSKDQSAILEETFKDHSTLNPKQKQALA
                     KQLGLRARQVEVWFQNRRARTKLKQTEVDCEFLRRCCENLTEENRRLQKEVTELRALK
                     LSPQFYMHMSPPTTLTMCPSCEHVSVPPPQPQAATSAHHRSLPVNAWAPATRISHGLT
                     FDALRPRS"
     exon            12153. .12312
                     /gene="dl4415w"
                     /number=1
     intron          12313. .12415
                     /gene="dl4415w"
                     /number=1
     exon            12416. .12714
                     /gene="dl4415w"
                     /number=2
     intron          12715. .12982
                     /gene="dl4415w"
                     /number=2
     exon            12983. .13062
                     /gene="dl4415w"
                     /number=3
     intron          13063. .13150
                     /gene="dl4415w"
                     /number=3
     exon            13151. .13466
                     /gene="dl4415w"
                     /number=4
     gene            14609. .16030
                     /gene="dl4420c"
     gene            complement(14609. .16030)
                     /gene="dl4420c"
     CDS             complement(14609. .16030)
                     /gene="dl4420c"
                     /note="sequence differences between cDNA and gDNA
                     similarity to hydroxyproline-rich glycoprotein (clone
                     Hyp3.6) - kidney bean (fragment)"
                     /codon_start=1
                     /product="glycoprotein homolog"
                     /protein_id="CAB46042.1"
                     /db_xref="GI:5302801"
                     /translation="MVEARSLKKPIQLGNKEDQNPRKFYSRFIFKALILTVLCAVVPV
                     FLSQTPELANQTRLLELLHLVFVGIAVSYGLFSRRNYDGGGGGGTSNSDHNKADHSNN
                     NSHSYVPKILEVSSVFNVGHESESEPSDDSSGDQRKFQTWKNKYHMKIPEVETRFVDR
                     VSSENREKPLLLPVRSLNYSRVSDSSGDNSGRWEKVRSKRELLKTLGDDNSDVLPSPI
                     PWRSRSSSSSSSSSKEVESLPSVKNLTTVESQPLIKNLTPPSSFSSPRKSNPIPNLAS
                     EFHPSPPPPPPPPPLPAFYNSSSRKDHPGIYRVERRESSVHKTKFAGGEFHPPPPPP
                     PPPPVEYYKSPPTKFRLSNERRKSSEQKMKRNAPKKVWWSDPIVESKEQDTEKNDQRS
                     NLGSKAVEESENGEQRRGENEIHDEVEKKIVEEEGVSEINNGSDVDKKADEFIAKFRE
                     QIRLQRIESIKRSTNKISANSSR"
     exon            complement(14609. .16030)
                     /gene="dl4420c"
```

```
                    /number=1
     repeat_region  16148. .16170
                    /note="23bp T tandem repeat"
     misc_feature   17691. .19115
                    /note="CDNA1KMY"
     gene           17793. .19171
                    /gene="dl4425c"
     gene           complement(join(17793. .18046,18147. .18250,18348. .18399,
                    18481. .18571,18886. .18960,19058. .19171))
                    /gene="dl4425c"
     CDS            complement(join(17793. .18046,18147. .18250,18348. .18399,
                    18481. .18571,18886. .18960,19058. .19171))
                    /gene="dl4425c"
                    /note="strong similarity to enoyl-CoA hydratase (EC
                    4.2.1.17), AU-specific RNA-binding protein, human"
                    /codon_start=1
                    /product="enoyl-CoA hydratase"
                    /protein_id="CAB10453.1"
                    /db_xref="GI:2245034"
                    /db_xref="SPTREMBL:O23520"
                    /translation="MLKSLQNAFESIHQDNSARVVMIRSLVPGVFCAGADLKERRTMS
                    PSEVHTYVNSLRYMFSFIEALSIPTIAAIEGAALGGGLEMALACDLRICGENAVFGLP
                    ETGLAIIPGAGGTQRLSRLVGRSVSKELIFTGRKIDAIEAANKGLVNICVTAGEAHEK
                    AIEMAQQINEKGPLAIKMAKKAIDEGIETNMASGLEVEEMCYQKLLNTQDRLEGLAAF
                    AEKRKPLYTGN"
     exon           complement(17793. .18046)
                    /gene="dl4425c"
                    /number=1
     intron         complement(18047. .18146)
                    /gene="dl4425c"
                    /number=1
     exon           complement(18147. .18250)
                    /gene="dl4425c"
                    /number=2
     intron         complement(18251. .18347)
                    /gene="dl4425c"
                    /number=2
     exon           complement(18348. .18399)
                    /gene="dl4425c"
                    /number=3
     intron         complement(18400. .18480)
                    /gene="dl4425c"
                    /number=3
     exon           complement(18481. .18571)
                    /gene="dl4425c"
                    /number=4
     intron         complement(18572. .18885)
                    /gene="dl4425c"
                    /number=4
     exon           complement(18886. .18960)
                    /gene="dl4425c"
                    /number=5
     intron         complement(18961. .19057)
                    /gene="dl4425c"
                    /number=5
     exon           complement(19058. .19171)
                    /gene="dl4425c"
                    /number=6
     gene           20179. .25115

  Query Match            18.5%;  Score 319.4;  DB 8;  Length 201471;
  Best Local Similarity  92.7%;  Pred. No. 3.1e-60;
  Matches  370;  Conservative    0;  Mismatches   21;  Indels    8;  Gaps    3;

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42342 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 42401

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||||| |||||     |||
Db      42402 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 42455

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      42456 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 42515

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db      42516 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 42575

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||| |||||||||||||||||  |||||| ||  |||||||||| ||||||||||
Db      42576 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 42634

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAG 1666
              |||||||||||||||||| |||||||||||||||||||||||  |||||| ||||||
Db      42635 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 42694

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||  ||
Db      42695 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 42733

Search completed: June 19, 2003, 21:25:47
Job time : 4500 secs
```

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:43:25 ; Search time 422 Seconds
(without alignments)
9189.434 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1722	100.0	1722	21	AAA63670	Nucleotide sequenc
	2	1722	100.0	1722	21	AAA47751	VRN2 nucleic acid.
	3	1569.6	91.1	1715	21	AAA63671	Nucleotide sequenc
	4	1569.6	91.1	1715	21	AAA47753	VRN2 nucleic acid.
	5	1537.6	89.3	1737	21	AAA47755	Aberrantly spliced
	6	979.8	56.9	1497	21	AAC51581	Arabidopsis thalia
	7	412.2	23.9	6338	21	AAA47752	VRN2 nucleic acid.
	8	319.4	18.5	5895	21	AAA47754	VRN2 nucleic acid.
	9	314.2	18.2	319	21	AAC51582	Arabidopsis thalia
	10	306.4	17.8	640	21	AAC44147	Arabidopsis thalia
	11	291.4	16.9	2280	21	AAZ36947	cDNA encoding a MP
	12	252.4	14.7	2248	21	AAZ36953	cDNA encoding an O
	13	230.8	13.4	558	21	AAA47758	At Hyp 2245035 (AT
	14	170.8	9.9	525	21	AAC54919	Arabidopsis thalia
	15	163.8	9.5	325	21	AAA47757	AI163743 Expressed
	16	126.6	7.4	463	21	AAA47756	Rice expressed seq
	17	71	4.1	8045	21	AAA65171	Neurospora crassa
c	18	69.6	4.0	567	21	AAA29550	HIV codon altered
	19	69	4.0	198	17	AAT17204	DNA-spanner oligon
	20	69	4.0	198	17	AAT17205	DNA-spanner oligon
	21	65.8	3.8	198	17	AAT17202	DNA-spanner oligon
	22	65.8	3.8	198	17	AAT17203	DNA-spanner oligon
	23	63.2	3.7	1216	21	AAZ36962	DNA fragment encod
	24	62.2	3.6	5580	21	AAZ36948	DNA encoding a MPC
	25	62	3.6	5511	21	AAA61847	Cryptosporidium pa
	26	62	3.6	5511	24	ABT04776	C parvum GP900 gen
	27	62	3.6	7334	21	AAA61846	Cryptosporidium pa
	28	62	3.6	7334	24	ABT04775	C parvum GP900 gen
	29	60.4	3.5	5163	19	AAV20700	Cryptosporidium pa
	30	60.4	3.5	5163	21	AAA61849	ORF encoding a por
	31	60.4	3.5	5163	24	ABT04778	C parvum GP900 gen
	32	60.4	3.5	5318	19	AAV20701	Cryptosporidium pa
	33	60.4	3.5	5318	21	AAA61848	DNA encoding a por
	34	60.4	3.5	5318	24	ABT04777	C parvum GP900 gen
c	35	59	3.4	61020	22	AAS46788	Tumour suppressor
c	36	56.8	3.3	2336	23	ABL25662	Drosophila melanog
	37	55.8	3.2	3300	21	AAZ55699	DNA encoding yeast
	38	54.6	3.2	4829	23	ABL13399	Drosophila melanog
c	39	54.6	3.2	39746	23	ABL13398	Drosophila melanog
c	40	54.2	3.1	8524	22	AAS46588	Tumour suppressor
c	41	54.2	3.1	8524	24	ABL33841	Human immune syste
	42	54	3.1	144	12	AAQ12515	CSP-2 peptide from
	43	53.4	3.1	1690	16	AAQ92524	Nicotiana alata ar
	44	53.2	3.1	2892	24	ABN79885	Fungal ZBC gene se
c	45	52.2	3.0	8662	24	ABL34637	Human metastasis a

ALIGNMENTS

RESULT 1
AAA63670
ID AAA63670 standard; cDNA; 1722 BP.
XX
AC AAA63670;
XX
DT 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
XX
KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; VRN2; ss.
XX
OS Arabidopsis sp.
XX
PN WO200046358-A2.
XX
PD 10-AUG-2000.
XX
PF 25-JAN-2000; 2000WO-GB00197.
XX
PR 05-FEB-1999; 99GB-0002660.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Johanson U, West J, Dean C;
XX
DR WPI; 2000-532899/48.
XX
PT New nucleic acid derived from the FRI locus of a plant, e.g.
PT Arabidopsis, encoding a polypeptide capable of specifically altering
PT the flowering time of a plant -

```
XX
PS   Example 5; Page 53-54; 73pp; English.
XX
CC   The present sequence represents a VRN2 sequence from variety Landsberg
CC   erecta. The specification describes a sequence which encodes a
CC   polypeptide capable of specifically altering the flowering time of a
CC   plant. The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
CC   of Arabidopsis. The FRI polynucleotide is used to transform plants, so
CC   that the flowering time of a plant is altered. This is used, for example,
CC   for plants in which the leaves or tubers are a commercial product, where
CC   it is desirable to avoid 'blotting' (initiation of flowers and stem
CC   elongation) at too early a stage. Conversely, it may be desirable to
CC   alter flowering under certain circumstances e.g. to vary flower
CC   production across the seasons.
XX
SQ   Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

  Query Match             100.0%;  Score 1722;  DB 21;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722
              |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 2
AAA47751
ID   AAA47751 standard; cDNA; 1722 BP.
XX
AC   AAA47751;
XX
DT   16-NOV-2000  (first entry)
XX
```

```
DE   VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana var 'Landsberg erecta'.
XX
FH   Key             Location/Qualifiers
FT   CDS             231..1568
FT                   /*tag= a
FT                   /product= "VRN2 polypeptide"
FT   mutation        replace(1199,A)
FT                   /*tag= b
FT                   /note= "found in mutant vin2-1; alters a Trp codon
FT                         to a STOP codon leading to a truncation of
FT                         the protein"
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00060.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 4; Fig 6; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

  Query Match             100.0%;  Score 1722;  DB 21;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy       61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy      121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy      181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy      241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy      301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy      361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy      421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy      481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy      541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy      601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy      661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy      721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy      781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy      841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy      901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy      961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy     1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy     1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy     1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy     1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy     1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy     1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy     1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
                 |||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
```

RESULT 3
AAA63671
ID AAA63671 standard; cDNA; 1715 BP.
XX
AC AAA63671;
XX
DT 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
XX
KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; VRN2; ss.
XX
OS Arabidopsis sp.
XX
PN WO200046358-A2.
XX
PD 10-AUG-2000.
XX
PF 25-JAN-2000; 2000WO-GB00197.
XX
PR 05-FEB-1999; 99GB-0002660.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Johanson U, West J, Dean C;
XX
DR WPI; 2000-532899/48.
XX
PT New nucleic acid derived from the FRI locus of a plant, e.g.
PT Arabidopsis, encoding a polypeptide capable of specifically altering
PT the flowering time of a plant -
XX
PS Example 5; Page 54-55; 73pp; English.
XX
CC The present sequence represents a VRN2 sequence from variety Columbia.
CC The specification describes a sequence which encodes a polypeptide
CC capable of specifically altering the flowering time of a plant.
CC The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
CC of Arabidopsis. The FRI polynucleotide is used to transform plants, so
CC that the flowering time of a plant is altered. This is used, for example,
CC for plants in which the leaves or tubers are a commercial product, where
CC it is desirable to avoid 'blotting' (initiation of flowers and stem
CC elongation) at too early a stage. Conversely, it may be desirable to
CC alter flowering under certain circumstances e.g. to vary flower
CC production across the seasons.
XX
SQ Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;

Query Match 91.1%; Score 1569.6; DB 21; Length 1715;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

```
Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              |||||||||||||||||||||||  ||||||||||||| ||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||| ||||| |||||||||| |||| ||||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              ||||||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718
              |||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720

Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838
              |||| |||||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900

Qy        899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958
              |||||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy        959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy       1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078
```

```
Db       1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy       1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138
Db       1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy       1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198
Db       1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy       1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258
Db       1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy       1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318
Db       1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy       1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378
Db       1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380

Qy       1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
Db       1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy       1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498
Db       1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy       1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558
Db       1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy       1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618
Db       1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy       1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
Db       1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAT 1705
Db       1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701
```

```
RESULT 4
AAA47753
ID   AAA47753 standard; cDNA; 1715 BP.
XX
AC   AAA47753;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana var 'Columbia'.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00061.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 5; Page 71; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;
```

```
  Query Match             91.1%;  Score 1569.6;  DB 21;  Length 1715;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   10;  Gaps    4;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598
Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658
Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718
Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720
```

```
Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838
              |||| ||||||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900

Qy        899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958
              |||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy        959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy       1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy       1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy       1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy       1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy       1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318
              || ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy       1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380

Qy       1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
              |||||||||||||||||||||||||||||| |||||       ||| | |||||||||
Db       1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy       1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy       1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558
              |||||||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db       1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy       1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618
              |||||||||||||| |||||| ||  |||||||||| ||||||||||||||||||||||||
Db       1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy       1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
              |||||||| |||||||||||||||||||||||  |||||| ||||||||||||||||||||
Db       1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||| ||
Db       1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701


RESULT 5
AAA47755
ID   AAA47755 standard; cDNA; 1737 BP.
XX
AC   AAA47755;
XX
DT   16-NOV-2000  (first entry)
XX
DE   Aberrantly spliced VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana var 'Columbia'.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00062.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 8; Page 74-75; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant. This cDNA was generated from an aberrantly spliced VRN2
CC   gene.
XX
SQ   Sequence 1737 BP; 543 A; 330 C; 359 G; 505 T; 0 other;

  Query Match            89.3%;  Score 1537.6;  DB 21;  Length 1737;
  Best Local Similarity  95.3%;  Pred. No. 0;
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   32;  Gaps    5;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              ||||||||||||||||||||||  |||||||||||| ||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||||| ||||| |||||||||| ||| ||||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360
```

```
Qy       359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
Db       361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy       419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
Db       421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy       479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCA------- 531
Db       481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGTGGG 540

Qy       532 ---------------AGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 576
Db       541 CAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 600

Qy       577 TTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 636
Db       601 TTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 660

Qy       637 TCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 696
Db       661 TCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 720

Qy       697 AACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 756
Db       721 AACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 780

Qy       757 TTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTA 816
Db       781 TTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGA 840

Qy       817 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 876
Db       841 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 900

Qy       877 TGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 936
Db       901 TGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 960

Qy       937 TATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 996
Db       961 TATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 1020

Qy       997 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1056
Db      1021 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1080

Qy      1057 AGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACG 1116
Db      1081 AGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACG 1140

Qy      1117 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1176
Db      1141 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1200

Qy      1177 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1236
Db      1201 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1260

Qy      1237 GTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTT 1296
Db      1261 GTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTT 1320

Qy      1297 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1356
Db      1321 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1380

Qy      1357 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAG 1416
Db      1381 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG 1440

Qy      1417 ACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACA 1476

Db      1441 ------TCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACA 1494

Qy      1477 ACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGG 1536
Db      1495 ACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGG 1554

Qy      1537 ACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGAT 1596
Db      1555 ACAATAACAGCAGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGAT 1613

Qy      1597 TTATCGGATTGTAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTT 1655
Db      1614 CTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATT 1673

Qy      1656 TTATATGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
Db      1674 TTATTTGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 1723
```

RESULT 6
AAC51581
ID AAC51581 standard; DNA; 1497 BP.
XX
AC AAC51581;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69038.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.9%; Score 979.8; DB 21; Length 1497;
Best Local Similarity 89.7%; Pred. No. 1.3e-232;
Matches 1118; Conservative 0; Mismatches 27; Indels 102; Gaps 2;

```
Qy       355 ACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGT 414
Db       320 AGCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGT 379
Qy       415 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAA 474
Db       380 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAA 439
Qy       475 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 534
Db       440 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 499
Qy       535 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCG 594
Db       500 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGTACATCT 559
Qy       595 AAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAG 654
Db       560 CTGTAAATCCAACCATAA---------------------------------------- 577
Qy       655 GAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGAC 714
Db       578 --------------------------TGTTTTGCAGCTCGAAACCTCGTAAGCGGAGAC 610
Qy       715 AAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCAC 774
Db       611 AAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCAC 670
Qy       775 CCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTT 834
Db       671 CCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTT 730
Qy       835 TAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCAC 894
Db       731 TAGGATATCCCGAGGCAACAGAACTTGCTGGACAATTTGAGATGACCAGCAACATTCCAC 790
Qy       895 CAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTG 954
Db       791 CAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTG 850
Qy       955 TGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGA------- 1007
Db       851 TGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGTACTG 910
Qy       1008 ------------------------AGCCACCTACTTCTTCAGAAACGCCAATTCTATC 1041
Db       911 CCATATTAACCGAGAAAATTTCTTCCAGCCACCTACTTCTTCAGAAACGCCAATTCTATC 970
Qy       1042 ATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGG 1101
Db       971 ATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGG 1030
Qy       1102 ATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGG 1161
Db       1031 ATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGG 1090
Qy       1162 ATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAA 1221
Db       1091 ATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAA 1150
Qy       1222 GGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGA 1281
Db       1151 GGGTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGAGA 1210
Qy       1282 AAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTAT 1341
Db       1211 AAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTAT 1270
Qy       1342 GGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATT 1401
Db       1271 GGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATT 1330
Qy       1402 GCCGTAATAGCTCAGACACCACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTG 1461
Db       1331 GGCGTAATAGCTCAGACACCACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTG 1390
Qy       1462 ACTCAAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACA 1521
Db       1391 ACTCAAACACCAACAACAATAACATTGTGAGTCATCCCAATGACATAAACAACAAGAACA 1450
Qy       1522 ATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAG 1568
Db       1451 ATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAG 1497
```

RESULT 7
AAA47752
ID AAA47752 standard; DNA; 6338 BP.
XX
AC AAA47752;
XX
DT 16-NOV-2000 (first entry)
XX
DE VRN2 nucleic acid.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time;
KW leaf size; leaf shape; shade avoidance response; reproduction;
KW breeding; pollination; cultivation; ds.
XX
OS Arabidopsis thaliana var 'Landsberg erecta'.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Claim 9; Page 68-70; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant.
XX
SQ Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;

Query Match 23.9%; Score 412.2; DB 21; Length 6338;
Best Local Similarity 98.8%; Pred. No. 9.4e-92;
Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1304 ACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363
Db       5102 AATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 5161
```

```
Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5162 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 5221

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5222 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 5281

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACACCAACAACAATAA 5281
             ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5282 CATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5342 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603

Qy       1604 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5402 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663

Qy       1664 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5462 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 5520
```

RESULT 8
AAA47754
ID AAA47754 standard; DNA; 5895 BP.
XX
AC AAA47754;
XX
DT 16-NOV-2000 (first entry)
XX
DE VRN2 nucleic acid.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time;
KW leaf size; leaf shape; shade avoidance response; reproduction;
KW breeding; pollination; cultivation; ds.
XX
OS Arabidopsis thaliana var 'Columbia'.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Claim 9; Page 72-74; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant.
XX
SQ Sequence 5895 BP; 1742 A; 1027 C; 1070 G; 2056 T; 0 other;

Query Match 18.5%; Score 319.4; DB 21; Length 5895;

Best Local Similarity 92.7%; Pred. No. 8.9e-69;
Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

```
Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
             | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4677 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 4736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4737 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 4790

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 4850
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4791 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 4850

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4851 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 4910

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
             |||||| ||||||||||||||||||  |||||| || |||||||||| ||||||||||||||
Db       4911 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 4969

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAG 1666
             |||||||||||||||||||| ||||||||||||||||||||||||  |||||| ||||||
Db       4970 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 5029

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
             |||||||||||||||||||||||||||||||||||||||  ||
Db       5030 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 5068
```

RESULT 9
AAC51582
ID AAC51582 standard; DNA; 319 BP.
XX
AC AAC51582;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69042.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.2%; Score 314.2; DB 21; Length 319;
Best Local Similarity 99.1%; Pred. No. 5.4e-68;
Matches 316; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1375 ACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACA 1434
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 ACAACTGCAATACCATCCTCGAGAATTGGCGTAATAGCTCAGACACCACCACCACCAACA 60

Qy       1435 ACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATC 1494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db         61 ACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGAGTC 120

Qy       1495 ATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACA 1554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACA 180

Qy       1555 AAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTA 1614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTA 240

Qy       1615 TTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCA 1674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCA 300

Qy       1675 ACTGATTACATTTTTAGTT 1693
             ||||||||||||||||||
Db        301 ACTGATTACATTTTTAGTT 319
```

RESULT 10
AAC44147
ID AAC44147 standard; DNA; 640 BP.
XX
AC AAC44147;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41806.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 17.8%; Score 306.4; DB 21; Length 640;
Best Local Similarity 98.8%; Pred. No. 6.1e-66;
Matches 318; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

```
Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 378

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 438

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAG- 1551
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 498

Qy       1552 ACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAAC 1611
             |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        499 ACAAAGTAATTAAATAGGAAAATCTCCGGCTTAGATGATACCGATTTATCGGATTGTAAC 558

Qy       1612 TTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTAT 1671
             ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        559 TTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAMCAAATTTTTTATATGTTAGTGTAT 618

Qy       1672 TCAACTGATTACATTTTTAGTT 1693
             |||||||||||||||||||||
Db        619 TCAACTGATTACATTTTTAGTT 640
```

RESULT 11
AAZ36947
ID AAZ36947 standard; cDNA; 2280 BP.
XX
AC AAZ36947;
XX
DT 13-MAR-2000 (first entry)
XX
DE cDNA encoding a MPC1 protein having flowering regulating activity.
XX
KW Flowering regulating activity; MPC1; flowering; germination;
KW super early flowering mutation; altered flowering time;
KW flowering regulating gene; food crop; vegetable; flowering inhibition;
KW productivity; ss.
XX

```
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             310..2145
FT                   /*tag=  a
FT                   /product= "flowering regulating protein"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;    99EP-0305077.
XX
PR   26-JUN-1998;    98JP-0180065.
PR   24-JUN-1999;    99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   P-PSDB; AAY53932.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 6; Page 16-21; 53pp; English.
XX
CC   The present sequence encodes a protein having a flowering regulating
CC   activity, which is designated MPC1. The genomic sequence is given
CC   in AAZ36948. A rice MPC1 is also disclosed in the specification. The
CC   rice and Arabidopsis cDNAs show significant homology with each other.
CC   A naturally occurring mutation of the MPC1 gene eliminates normal
CC   flowering regulating ability of plants, and leads to flowering
CC   immediately after germination (super early flowering mutation). The
CC   MPC1 polynucleotide sequence can be used to produce plants with altered
CC   flowering times in comparison with wild type plants, by enhancing
CC   or inhibiting the expression of the flowering regulating gene. Antisense
CC   polynucleotides can be used to reproduce the effects of the mutated MPC1
CC   gene. This alteration can be used to increase the yield of food crops.
CC   Flowering inhibition of vegetables increases their productivity.
XX
SQ   Sequence 2280 BP; 663 A; 445 C; 521 G; 651 T; 0 other;

  Query Match              16.9%;  Score 291.4;  DB 21;  Length 2280;
  Best Local Similarity    56.6%;  Pred. No. 5.2e-62;
  Matches  563;  Conservative    0;  Mismatches  426;  Indels    6;  Gaps    1;

Qy        423 ACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTT 482
Db       1150 ACAGGAAATGTAGTTTTCAACTACAGATACTATAACAACAAATTGCAGAAGACTGAAGTA 1209
Qy        483 AGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAA 542
Db       1210 ACTGAAGACTTTTCTTGTCCATTCTGCTTAGTAAAATGTGCCAGTTTCAAGGGCCTGAGA 1269
Qy        543 TTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATAC 602
Db       1270 TATCACTTGCCATCAACCCACGATCTCCTCAATTTCGAGTTTTGGGTAACTGAAGAATTT 1329
Qy        603 CAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGAT 662
Db       1330 CAGGCGGTAAATGTCTCCCTCAAGACTGAGACAATGATATCCAAGGTTAATGAGGATGAC 1389
Qy        663 GACGATAAATTTGAGC------CCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAA 716
Db       1390 GTTGACCCAAAGCAGCAAACTTTCTTTTTTTCTTCCAAAAAATTCAGACGGAGGAGGCAA 1449
Qy        717 AGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCC 776
Db       1450 AAGAGTCAGGTACGGAGCTCAAGGCAAGGGCCTCATCTTGGATTAGGTTGCGAGGTGCTA 1509
Qy        777 AGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTA 836

Db       1510 GATAAGACTGATGATGCTCATTCTGTTAGAAGTGAGAAGAGCCGAATACCACCTGGAAAG 1569
Qy        837 GGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCA 896
Db       1570 CATTACGAAAGAATTGGGGGTGCTGAGTCTGGTCAAAGAGTTCCTCCTGGCACGAGTCCT 1629
Qy        897 GCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTG 956
Db       1630 GCAGACGTGCAATCATGTGGGGATCCAGATTATGTGCAGTCGATAGCTGGAAGTACAATG 1689
Qy        957 GTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTA 1016
Db       1690 TTGCAGTTTGCAAAAACGAGGAAAATATCTATAGAACGGTCGGACTTGAGGAACCGAAGC 1749
Qy       1017 CTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAA 1076
Db       1750 CTCCTTCAGAAGAGACAGTTGTTCCACTCTCATCGAGCTCAGCCCATGGCTCTAGAACAA 1809
Qy       1077 GTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGAT 1136
Db       1810 GTACTTTCGGACCGGGATAGTGAAGATGAAGTTGATGATGATGTGGCAGATTTTGAAGAT 1869
Qy       1137 CGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTT 1196
Db       1870 AGAAGGATGCTCGATGATTTCGTTGATGTGACTAAAGATGAGAAACAGATGATGCACATG 1929
Qy       1197 TGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGT 1256
Db       1930 TGGAACTCGTTTGTGAGGAAGCAGCGAGTATTAGCAGATGGTCACATTCCATGGGCATGC 1989
Qy       1257 GAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGT 1316
Db       1990 GAGGCATTCTCAAGATTGCACGGACCCATCATGGTTCGAACACCGCACTTGATTTGGTGC 2049
Qy       1317 TGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAAC 1376
Db       2050 TGGAGAGTGTTTATGGTGAAACTGTGGAACCACGGTCTTCTTGATGCCCGAACCATGAAC 2109
Qy       1377 AACTGCAATACCATCCTCGAGAATTGCCGTAATAG 1411
Db       2110 AACTGTAATACCTTTCTCGAACAGCTCCAAATTTG 2144

RESULT 12
AAZ36953
ID   AAZ36953 standard; cDNA; 2248 BP.
XX
AC   AAZ36953;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA encoding an Os-MPC1 protein having flowering regulating activity.
XX
KW   Flowering regulating activity; MPC1; flowering; germination;
KW   super early flowering mutation; altered flowering time;
KW   flowering regulating gene; food crop; vegetable; flowering inhibition;
KW   productivity; ss.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   CDS             86..1900
FT                   /*tag=  a
FT                   /product= "flowering regulating protein"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;    99EP-0305077.
XX
PR   26-JUN-1998;    98JP-0180065.
```

```
PR   24-JUN-1999;   99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   P-PSDB; AAY53933.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 6; Page 40-44; 53pp; English.
XX
CC   The present sequence encodes a protein having a flowering regulating
CC   activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
CC   disclosed in the specification. The rice and Arabidopsis cDNAs show
CC   significant homology with each other. A naturally occurring mutation
CC   of the MPC1 gene eliminates normal flowering regulating ability of
CC   plants, and leads to flowering immediately after germination (super
CC   early flowering mutation). The MPC1 polynucleotide sequence can be
CC   used to produce plants with altered flowering times in comparison
CC   with wild type plants, by enhancing or inhibiting the expression of
CC   the flowering regulating gene. Antisense polynucleotides can be used
CC   to reproduce the effects of the mutated MPC1 gene. This alteration
CC   can be used to increase the yield of food crops. Flowering inhibition
CC   of vegetables increases their productivity.
XX
SQ   Sequence 2248 BP; 641 A; 454 C; 493 G; 660 T; 0 other;

  Query Match            14.7%;  Score 252.4;  DB 21;  Length 2248;
  Best Local Similarity  56.2%;  Pred. No. 2.4e-52;
  Matches  555;  Conservative    0;  Mismatches  376;  Indels   57;  Gaps    2;

Qy       424 CTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTA 483
Db       939 CTGGCAATGTGCTTTTTAACTACAAATACTACAATAATACTATGCAAAAAACCGAAGTCA 998

Qy       484 GGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAAT 543
Db       999 CTGAAGATTTTTCTTGCCCATTTTGCTTGGTACCATGTGGCAGCTTTAAGGGTCTAGGAT 1058

Qy       544 TTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACC 603
Db      1059 GTCACCTAAACGCATCGCATGACCTTTTCCATTATGAGTTTTGGATATCTGAAGAGTGCC 1118

Qy       604 AGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG---GAAGAAGGAAGTG 660
Db      1119 AGGCTGTTAATGTTAGTCTGAAGACTGATTCTTGGAGAACAGAGCTTTTGGCTGAGGGAG 1178

Qy       661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
Db      1179 TTGATCCAAGACATCAAACATTTTCGTACCGCTCAAGATTTAAGAAGCGTAAAAGGGTGG 1238

Qy       721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
Db      1239 AAATCTCAAGTGATAAAATTAGGCATGTACATCCACATATTGTGGATTCAGGATCACCTG 1298

Qy       781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
Db      1299 AAGATGCCCAGGCAGGATCTGAAGACGATTACGTGCAGAGGGAAAATGGTAGTTCTGTAG 1358

Qy       841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
Db      1359 CACACGCTTCTGTTGATCCTGCTAATTCATTACACGGTAGCAATCTTTCAGCACC----- 1413

Qy       901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
Db      1414 ----------------------------------------------AACAGTGTTAC 1424

Qy       961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
Db      1425 AGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGAAATCGGCAGCTCC 1484

Qy      1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
Db      1485 TACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCATGGAGCAAAGTTT 1544

Qy      1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
Db      1545 TCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGATTTTGAAGATAGAA 1604

Qy      1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
Db      1605 GAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATTATGCATATGTGGA 1664

Qy      1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
Db      1665 ATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCCTGGGCATGCGAAG 1724

Qy      1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
Db      1725 CATTCTCGCAGTTTCATGGACAAGAACTTGTACAAAATCCAGCTCTACTATGGTGTTGGA 1784

Qy      1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
Db      1785 GGTTTTTTATGGTCAAACTCTGGAACCACAGTCTACTGGATGCGCGAGCCATGAATGCCT 1844

Qy      1381 GCAATACCATCCTCGAGAATTGCCGTAA 1408
Db      1845 GCAACACAATTCTTGAAGGCTACCTGAA 1872


RESULT 13
AAA47758
ID   AAA47758 standard; cDNA; 558 BP.
XX
AC   AAA47758;
XX
DT   16-NOV-2000  (first entry)
XX
DE   At Hyp 2245035 (ATFCA7_4) (modified) cDNA.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00065.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Disclosure; Page 76; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
```

```
CC   the plant. This cDNA represents another Arabidopsis thaliana sequence
CC   having homology to the VRN2 gene.
XX
SQ   Sequence 558 BP; 191 A; 122 C; 112 G; 133 T; 0 other;

  Query Match             13.4%;  Score 230.8;  DB 21;  Length 558;
  Best Local Similarity   69.4%;  Pred. No. 3.1e-47;
  Matches  361;  Conservative    0;  Mismatches  147;  Indels   12;  Gaps    3;

Qy        978 AAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCC---ACCTACTTCTTCAGAAACGCCAA 1034
              ||| |  ||||| |||||||| |||||| |||||    || | | |||||| ||||||||
Db         31 AAGGTCCCTGCTAAGCGATCAAAGGCTACAAGCCACTACTTGCCTCTTCATAAACGCCAG 90

Qy       1035 TTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGAT 1094
              |||||||||||| |  |  | |||||| || | |||||||||||| |||||||||||| |||
Db         91 TTCTATCATTCCCGAACCGGTCAGCCATTGTCACTTGAGCAAGTTATGTCTGACCGAGAT 150

Qy       1095 AGCGAGGATGAAGTCG------ATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTT 1148
              |||||  |||| ||||      |||| |||| |||| || | |||||  ||||||||||
Db        151 AGCGAAAATGACGTCGACAAAAATGATGATGCTGCACATCTCGAAGAAAGCCAGATGCTT 210

Qy       1149 GATGACTTTGTGGATGTGAATAAA---GATGAAAAGCAATTCATGCATCTTTGGAACTCG 1205
              |||  |   |||||| |||| ||    |  | | ||  ||||||  | ||||||||||
Db        211 AATGGTTCCATGGATGAGAATGAAATCGTAGCAGAGAGATTCATAAAACTTTGGAACTCC 270

Qy       1206 TTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTT 1265
              ||||| | | | |||||| || | ||||||| ||||||  ||||||||||||||||||
Db        271 TTTGTTAAACAGCAAAGGATTGTTGCAGATGCTCATATTCCTTGGGCATGTGAAGCATTC 330

Qy       1266 TCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTG 1325
              ||||||||  ||  | ||||| ||| | || ||| |||||||  || ||| ||||||
Db        331 TCAAGATTACACCTGCAAGAGCTGCGCAGTAACTTATCACTCGACTTGTGCTGGAGACAA 390

Qy       1326 TTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAAT 1385
              ||  |||| |||| |||| |  |||||||| | |||  || |||||| ||||| ||||||
Db        391 TTCATGATCAAACAATGGGATTATGGACTTCTTGACAGAGTCACCATGAACAAATGCAAT 450

Qy       1386 ACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGT 1445
              |||||| || |  ||          |   ||     ||  | ||| |||||||  ||
Db        451 ACCATCATCTACCATAATATCTCAACTACCAACGATGACATAAACAATAACAACACAAGG 510

Qy       1446 GTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACA 1485
              |  |  |   |  ||       |  | |  |||  ||||
Db        511 ACGACTGATAATATGGATGTTGTCGACGATGACATAAACA 550


RESULT 14
AAC54919
ID   AAC54919 standard; DNA; 525 BP.
XX
AC   AAC54919;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 79539.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
```

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.9%; Score 170.8; DB 21; Length 525;
Best Local Similarity 96.2%; Pred. No. 2.2e-32;
Matches 175; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy        355 ACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGT 414
              | |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db        326 AGCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGT 385

Qy        415 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAA 474
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| | ||
Db        386 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAA 445

Qy        475 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 534
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 505

Qy        535 GG 536
                |
Db        506 TG 507
```

RESULT 15
AAA47757
ID AAA47757 standard; cDNA; 325 BP.
XX
AC AAA47757;
XX
DT 16-NOV-2000 (first entry)
XX
DE AI163743 Expressed sequence tag (modified) cDNA.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time; poplar;
KW leaf size; leaf shape; shade avoidance response; reproduction; EST;
KW breeding; pollination; cultivation; expressed sequecne tag; ss.
XX
OS Populus tremula.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX

PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
DR P-PSDB; AAB00064.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 76; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA represents expressed sequence tag clone AI163743
CC from poplar trees (Populus tremula x Populus tremuloides) which has
CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
XX
SQ Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;

Query Match 9.5%; Score 163.8; DB 21; Length 325;
Best Local Similarity 76.4%; Pred. No. 9.8e-31;
Matches 201; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010
              ||  || | |    | | ||||||||||| |||||   ||| || || ||     |||| |
Db         62 GCCATGCTACAATTTGCAAAGACAAGAAAATTATCAATTGAACGGTCTGACATGAGAAAC 121

Qy       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070
              |    ||| ||||| ||||| ||||| | ||| || || ||||  ||||||||||   |
Db        122 CGTACACTCCTTCACAAACGACAATTTTTTCACTCACATAGAGCTCAGCCAATGGCAGCT 181

Qy       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130
              |||||||| ||||| || |||||||| |||||||||| || || ||||||||||||||
Db        182 GAGCAAGTTATGTCAGATCGGGATAGTGAGGATGAAGTTGACGATGATGTTGCAGATTTT 241

Qy       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190
              |||||  ||   |||||||||||| ||||| || |||| ||||||||| |||||| | |||
Db        242 GAAGACCGAAGGATGCTTGATGATTTTGTAGACGTGACTAAAGATGAGAAGCAAATGATG 301

Qy       1191 CATCTTTGGAACTCGTTTGTAAG 1213
              ||  | |||||||| ||||| ||
Db        302 CACTTGTGGAACTCATTTGTGAG 324
```

Search completed: June 19, 2003, 20:10:38
Job time : 430 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:56:48 ; Search time 90 Seconds
(without alignments)
5867.749 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	69	4.0	198	5	PCT-US95-10668-3	Sequence 3, Appli
	2	69	4.0	198	5	PCT-US95-10668-4	Sequence 4, Appli
	3	65.8	3.8	198	5	PCT-US95-10668-1	Sequence 1, Appli
	4	65.8	3.8	198	5	PCT-US95-10668-2	Sequence 2, Appli
	5	62	3.6	5511	3	US-08-928-361B-2	Sequence 2, Appli
	6	62	3.6	7334	3	US-08-928-361B-1	Sequence 1, Appli
	7	60.4	3.5	5163	3	US-08-700-651-1	Sequence 1, Appli
	8	60.4	3.5	5163	3	US-08-928-361B-4	Sequence 4, Appli
	9	60.4	3.5	5318	3	US-08-700-651-2	Sequence 2, Appli
	10	60.4	3.5	5318	3	US-08-928-361B-3	Sequence 3, Appli
	11	59.2	3.4	1086	1	US-08-415-751-27	Sequence 27, Appl
c	12	59.2	3.4	1086	1	US-08-415-751-28	Sequence 28, Appl
	13	53.4	3.1	1690	1	US-08-276-452A-24	Sequence 24, Appl
	14	53.4	3.1	1690	2	US-08-798-744-24	Sequence 24, Appl
	15	51	3.0	2032	4	US-09-241-581B-5	Sequence 5, Appli
	16	51	3.0	2032	4	US-08-265-428-5	Sequence 5, Appli
	17	51	3.0	2032	5	PCT-US95-07721-5	Sequence 5, Appli
c	18	49.8	2.9	1086	1	US-08-415-751-47	Sequence 47, Appl
	19	48.2	2.8	688	4	US-08-998-416-915	Sequence 915, App
	20	47.6	2.8	1430	1	US-08-276-452A-25	Sequence 25, Appl
	21	47.6	2.8	1430	2	US-08-798-744-25	Sequence 25, Appl
	22	45.8	2.7	2190	4	US-09-625-188-19	Sequence 19, Appl
	23	43.6	2.5	1107	2	US-08-991-300-1	Sequence 1, Appli
	24	43.4	2.5	3292	1	US-07-814-964-12	Sequence 12, Appl
	25	43.4	2.5	3292	1	US-08-258-442-12	Sequence 12, Appl
	26	43.4	2.5	3292	1	US-08-328-809-7	Sequence 7, Appli
	27	43.4	2.5	3292	5	PCT-US92-11107-12	Sequence 12, Appl
c	28	40.8	2.4	87350	3	US-08-781-891-79	Sequence 79, Appl
c	29	40.8	2.4	87543	4	US-09-791-211-3	Sequence 3, Appli
c	30	39.8	2.3	915	4	US-09-134-001C-1741	Sequence 1741, Ap
	31	39.6	2.3	1529	2	US-08-477-451-17	Sequence 17, Appl
c	32	39.6	2.3	1529	2	US-08-477-451-21	Sequence 21, Appl
	33	39	2.3	2295	1	US-08-375-300-3	Sequence 3, Appli
	34	39	2.3	2295	3	US-09-177-431-3	Sequence 3, Appli
	35	39	2.3	2295	5	PCT-US95-16930-3	Sequence 3, Appli
	36	39	2.3	4080	1	US-08-375-300-1	Sequence 1, Appli
	37	39	2.3	4080	3	US-09-177-431-1	Sequence 1, Appli
	38	39	2.3	4080	5	PCT-US95-16930-1	Sequence 1, Appli
c	39	38.8	2.3	658	4	US-08-998-416-595	Sequence 595, App
c	40	38.8	2.3	912	4	US-09-134-001C-1352	Sequence 1352, Ap
c	41	38.6	2.2	399	3	US-08-713-569-6	Sequence 6, Appli
	42	38.4	2.2	444	4	US-09-178-509-2	Sequence 2, Appli
c	43	38.4	2.2	450	1	US-08-090-523-28	Sequence 28, Appl
c	44	38.4	2.2	450	1	US-08-398-627-28	Sequence 28, Appl
c	45	38.4	2.2	450	1	US-08-406-857-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
; Sequence 3, Application PC/TUS9510668
; GENERAL INFORMATION:
; APPLICANT: James Eberwine
; TITLE OF INVENTION: A Method of Sequencing Proteins by
; TITLE OF INVENTION: Epitope Ordering and Protein
; TITLE OF INVENTION: Restriction Mapping
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10668
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/294,133
; FILING DATE: August 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: PENN-0137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
PCT-US95-10668-3

Query Match 4.0%; Score 69; DB 5; Length 198;
Best Local Similarity 60.3%; Pred. No. 1.2e-09;
Matches 114; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432

```
Db          1 CAACAACAACAACAACAACAACAACAACAACAACAAGAATTCCAACAACAACAACAACAACA 60

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||||      |       ||   ||   |||| |||||||| |||| |
Db         61 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 120

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                |   |||  |||  |||||||| |||||     |||||||  |||| |||| ||   |
Db        121 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 180

Qy       1553 CAAAGTAAT 1561
              |||  | ||
Db        181 CAAGCTGAT 189


RESULT 2
PCT-US95-10668-4
; Sequence 4, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-4

  Query Match             4.0%;  Score 69;  DB 5;  Length 198;
  Best Local Similarity   60.3%;  Pred. No. 1.2e-09;
  Matches  114;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              |||||||  ||| | || |  | | ||   |   || |  ||  ||| || || || |||
Db          1 CAACAACAACAACAACAACAACAACAACAACAACAAGAATTCCAACAACAACAACAACAA 60

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||||      |       ||   ||   |||| |||||||| |||| |
Db         61 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 120

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                |   |||  |||  |||||||| |||||     |||||||  |||| |||| ||   |
Db        121 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 180

Qy       1553 CAAAGTAAT 1561
              |||  | ||
Db        181 CAAGCTGAT 189


RESULT 3
PCT-US95-10668-1
; Sequence 1, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-1

  Query Match             3.8%;  Score 65.8;  DB 5;  Length 198;
  Best Local Similarity   59.3%;  Pred. No. 8.9e-09;
  Matches  112;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              |||||||  ||| | || |  | | ||   |   || |  ||  ||| || || || |||
Db          1 CAACAACAACAACAACAACAACAACAACAACAACAAGGATCCCAACAACAACAACAACAA 60

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||||      |       ||   ||   |||| |||||||| |||| |
Db         61 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 120

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                |   |||  |||  |||||||| |||||     |||||||  |||| |||| ||   |
Db        121 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 180

Qy       1553 CAAAGTAAT 1561
              |||  | ||
Db        181 CAAGCTGAT 189


RESULT 4
```

```
PCT-US95-10668-2
; Sequence 2, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-2

  Query Match             3.8%;  Score 65.8;  DB 5;  Length 198;
  Best Local Similarity   59.3%;  Pred. No. 8.9e-09;
  Matches  112;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              ||||||| ||| | || |  | | ||   |   ||     |  ||| || || || |||
Db          1 CAACAACAACAACAACAACAACAACAACAACAACAAGGATCCCAACAACAACAACAACAA 60

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||||      |      ||   ||   |||| |||||||| ||||      |
Db         61 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAA 120

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                |   |||  |||  |||||||| |||||     |||||||  |||| |||| ||   |
Db        121 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAA 180

Qy       1553 CAAAGTAAT 1561
              |||  | ||
Db        181 CAAGCTGAT 189


RESULT 5
US-08-928-361B-2
; Sequence 2, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-2

  Query Match             3.6%;  Score 62;  DB 3;  Length 5511;
  Best Local Similarity   57.1%;  Pred. No. 5.2e-07;
  Matches  113;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||| |      | |  ||| |   || | |||| |
Db       2150 CTACTAAGGCAACAACAACCACAACAATTCTTAATCCAATCATTACAACAACAACTCAAA 2209

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || |||||||||||  ||           |||   |     ||| ||||||
Db       2210 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 2269

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |   ||  |  |||  ||| ||||| |  |||||     |||||||  |
Db       2270 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 2329

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       2330 CAGTGCCAACGACAACTA 2347


RESULT 6
US-08-928-361B-1
; Sequence 1, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
```

```
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7334 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-1

  Query Match             3.6%;  Score 62;  DB 3;  Length 7334;
  Best Local Similarity   57.1%;  Pred. No. 6e-07;
  Matches  113;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||| |      | |  ||| |   || | |||| |
Db       3818 CTACTAAGGCAACAACAACCACAACAATTCTTAATCCAATCATTACAACAACAACTCAAA 3877

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || ||||||||||||  ||          |||   |     ||| ||||||
Db       3878 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 3937

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |    ||  |  |||  |||  ||||| |  |||||     |||||||  |
Db       3938 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 3997

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       3998 CAGTGCCAACGACAACTA 4015


RESULT 7
US-08-700-651-1
; Sequence 1, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 5163
;   TYPE: DNA
;   ORGANISM: Cryptosporidium parvum
US-08-700-651-1

  Query Match             3.5%;  Score 60.4;  DB 3;  Length 5163;
  Best Local Similarity   56.6%;  Pred. No. 1.4e-06;
  Matches  112;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||  |      | |  ||| |   || | |||| |
Db       1799 CTACTAAGGCAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAA 1858

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || ||||||||||||  ||          |||   |     ||| ||||||
Db       1859 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 1918

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |    ||  |  |||  |||  ||||| |  |||||     |||||||  |
Db       1919 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 1978

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       1979 CAGTGCCAACGACAACTA 1996


RESULT 8
US-08-928-361B-4
; Sequence 4, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
;   APPLICANT:  Petersen, Carolyn
;   TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;   TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;   TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;   TITLE OF INVENTION:  SPECIES INFECTIONS
;   NUMBER OF SEQUENCES:  30
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;     STREET:  385 Sherman Avenue, Suite 6
;     CITY:  Palo Alto
;     STATE:  CA
;     COUNTRY:  USA
;     ZIP:  94306-1840
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/928,361B
;     FILING DATE:  12-SEP-1997
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 60/026,062
;     FILING DATE:  13-SEP-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Verny, Hana
;     REGISTRATION NUMBER:  30,518
;     REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  650-324-1677
;     TELEFAX:  650-324-1678
; INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  5163 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-4

  Query Match             3.5%;  Score 60.4;  DB 3;  Length 5163;
  Best Local Similarity   56.6%;  Pred. No. 1.4e-06;
```

```
  Matches  112;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||  |       | |  ||| |   || | |||| |
Db       1799 CTACTAAGGCAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAA 1858

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || ||||||||||||  ||          |||   |      ||| ||||||
Db       1859 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 1918

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |   ||  |  |||  |||  ||||| |  |||||     |||||||  |
Db       1919 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 1978

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       1979 CAGTGCCAACGACAACTA 1996


RESULT 9
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 5318
;   TYPE: DNA
;   ORGANISM: Cryptosporidium parvum
US-08-700-651-2

  Query Match             3.5%;  Score 60.4;  DB 3;  Length 5318;
  Best Local Similarity   56.6%;  Pred. No. 1.4e-06;
  Matches  112;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||  |       | |  ||| |   || | |||| |
Db       1799 CTACTAAGGCAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAA 1858

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || ||||||||||||  ||          |||   |      ||| ||||||
Db       1859 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 1918

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |   ||  |  |||  |||  ||||| |  |||||     |||||||  |
Db       1919 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 1978

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       1979 CAGTGCCAACGACAACTA 1996


RESULT 10
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-3

  Query Match             3.5%;  Score 60.4;  DB 3;  Length 5318;
  Best Local Similarity   56.6%;  Pred. No. 1.4e-06;
  Matches  112;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||| || || || |||| ||  |||  |       | | .||| |   || | |||| |
Db       1799 CTACTAAGGCAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAA 1858

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              |||| || || ||||||||||||  ||          |||   |      ||| ||||||
Db       1859 AACCAACAACAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAA 1918

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              ||| |  |  |   ||  |  |||  |||  ||||| |  |||||     |||||||  |
Db       1919 CAACATTAAAGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAA 1978

Qy       1538 CAATAACAGCAGAGACAA 1555
              || |  || |    || |
Db       1979 CAGTGCCAACGACAACTA 1996


RESULT 11
US-08-415-751-27
; Sequence 27, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
```

```
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1086 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
US-08-415-751-27

  Query Match             3.4%;  Score 59.2;  DB 1;  Length 1086;
  Best Local Similarity   56.8%;  Pred. No. 1.3e-06;
  Matches  109;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy       1376 CAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCACCAACAA 1435
                 ||||  |||   || |  |    | | |  | ||  |    ||| | | |  ||||||||
Db        341 CAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAA 400

Qy       1436 CAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCA 1495
                 ||||||||       |  || ||||  |  | | ||  | | |||| ||||       | |
Db        401 CAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAA 460

Qy       1496 TCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAA 1555
                ||||      |||||||| |||||     |||||||  |||| |||| ||| | |||
Db        461 AGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAA 520

Qy       1556 AGTAATTAAATA 1567
                |  |   | ||
Db        521 CGACAACTACTA 532


RESULT 12
US-08-415-751-28/c
; Sequence 28, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1086 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
US-08-415-751-28

  Query Match             3.4%;  Score 59.2;  DB 1;  Length 1086;
  Best Local Similarity   56.8%;  Pred. No. 1.3e-06;
  Matches  109;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy       1376 CAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCACCAACAA 1435
                 ||||  |||   || |  |    | | |  | ||  |    ||| | | |  ||||||||
Db         746 CAACAACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAA 687

Qy       1436 CAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCA 1495
                 ||||||||       |  || ||||  |  | | ||  | | |||| ||||       | |
Db         686 CAACAACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAA 627

Qy       1496 TCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAA 1555
                ||||      |||||||| |||||     |||||||  |||| |||| ||| | |||
Db         626 AGCCAATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAA 567

Qy       1556 AGTAATTAAATA 1567
                |  |   | ||
Db         566 CGACAACTACTA 555


RESULT 13
US-08-276-452A-24
; Sequence 24, Application US/08276452A
; Patent No. 5646029
;  GENERAL INFORMATION:
;    APPLICANT:  Chen, Chao-Guang
;    APPLICANT:  Mau, Shaio-Lim
;    APPLICANT:  Du, He
```

```
;    APPLICANT:  Gane, Alison M
;    APPLICANT:  Bacic, Antony
;    APPLICANT:  Clarke, Adrienne E
;    TITLE OF INVENTION:  Plant Arabinogalactan Protein (AGP) Genes
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,452A
;      FILING DATE:  18-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Caruthers, Jennie M.
;      REGISTRATION NUMBER:  34,464
;      REFERENCE/DOCKET NUMBER:  27-91A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303)499-8080
;      TELEFAX:  (303)499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1690 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  60..1442
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..38
;      OTHER INFORMATION:  /note= "Nucleotide sequence
;      OTHER INFORMATION:  obtained by PCR which does not overlap with the
;      OTHER INFORMATION:  cDNA clone"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  60..128
;      OTHER INFORMATION:  /note= "Predicted transmembrane
;      OTHER INFORMATION:  segment"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Derived amino acid sequence
;      OTHER INFORMATION:  corresponding to the peptide sequence by protein
;      OTHER INFORMATION:  microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Amino acids 27 to 36, 38,
;      OTHER INFORMATION:  and 40 are identical to that in the peptide
;      OTHER INFORMATION:  obtained by direct microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Amino acid 26 may also be
;      OTHER INFORMATION:  Ala; 37 and 39 can also be undetermined residues"
US-08-276-452A-24

  Query Match             3.1%;  Score 53.4;  DB 1;  Length 1690;
  Best Local Similarity   55.7%;  Pred. No. 6.4e-05;
  Matches  102;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              | | || | ||  | || |  | |    | |    | || |   ||| || || || |||
Db        920 CCAGAATTACATGAACAACAACAATGGCTTCTCCGAGAGTTACAACAACAACAACAACAA 979

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||     | |    |    | |  |||| |||||| || || | |||
Db        980 CAACAACAACAACAACGTTTTCTCTGAGAATTACAACAACAACAATAACAATAATGTTTT 1039

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                  ||| |||  |||||||| |||||   ||    | | || ||| ||| ||   |
Db       1040 CTCCGAGAATTACAACAACAACAATAACAACAATGCTTTCTACGAGAATTACAACAACAA 1099

Qy       1553 CAA 1555
              |||
Db       1100 CAA 1102


RESULT 14
US-08-798-744-24
; Sequence 24, Application US/08798744
; Patent No. 5830747
;  GENERAL INFORMATION:
;    APPLICANT:  Chen, Chao-Guang
;    APPLICANT:  Mau, Shaio-Lim
;    APPLICANT:  Du, He
;    APPLICANT:  Gane, Alison M
;    APPLICANT:  Bacic, Antony
;    APPLICANT:  Clarke, Adrienne E
;    TITLE OF INVENTION:  Plant Arabinogalactan Protein (AGP) Genes
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/798,744
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/276,452
;      FILING DATE:  18-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Caruthers, Jennie M.
;      REGISTRATION NUMBER:  34,464
;      REFERENCE/DOCKET NUMBER:  27-91A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303)499-8080
;      TELEFAX:  (303)499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1690 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  60..1442
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..38
;      OTHER INFORMATION:  /note= "Nucleotide sequence
;      OTHER INFORMATION:  obtained by PCR which does not overlap with the
;      OTHER INFORMATION:  cDNA clone"
```

```
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  60..128
;      OTHER INFORMATION:  /note= "Predicted transmembrane
;      OTHER INFORMATION:  segment"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Derived amino acid sequence
;      OTHER INFORMATION:  corresponding to the peptide sequence by protein
;      OTHER INFORMATION:  microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Amino acids 27 to 36, 38,
;      OTHER INFORMATION:  and 40 are identical to that in the peptide
;      OTHER INFORMATION:  obtained by direct microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  135..179
;      OTHER INFORMATION:  /note= "Amino acid 26 may also be
;      OTHER INFORMATION:  Ala; 37 and 39 can also be undetermined residues"
US-08-798-744-24

  Query Match             3.1%;  Score 53.4;  DB 2;  Length 1690;
  Best Local Similarity   55.7%;  Pred. No. 6.4e-05;
  Matches  102;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              | | || | ||  | || |  | |    | |    | || |   ||| || || || |||
Db        920 CCAGAATTACATGAACAACAACAATGGCTTCTCCGAGAGTTACAACAACAACAACAACAA 979

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||     | |    |     | |  |||| |||||| || || | |||
Db        980 CAACAACAACAACAACGTTTTCTCTGAGAATTACAACAACAACAATAACAATAATGTTTT 1039

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                  ||| |||  |||||||| |||||   ||    | | || ||| ||| ||   |
Db       1040 CTCCGAGAATTACAACAACAACAATAACAACAATGCTTTCTACGAGAATTACAACAACAA 1099

Qy       1553 CAA 1555
              |||
Db       1100 CAA 1102


RESULT 15
US-09-241-581B-5
; Sequence 5, Application US/09241581B
; Patent No. 6350859
;   GENERAL INFORMATION:
;        APPLICANT: Massachusetts Institute of Technology
;        TITLE OF INVENTION: Class BI and CI Scavenger Receptors
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Patrea L. Pabst
;             STREET: 2800 One Atlantic Center
;                     1201 West Peachtree Street
;             CITY: Atlanta
;             STATE: Georgia
;             COUNTRY: USA
;             ZIP: 30309-3450
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/241,581B
;             FILING DATE: 02-Feb-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pabst, Patrea L.
;             REGISTRATION NUMBER: 31,284
;             REFERENCE/DOCKET NUMBER: MIT6620
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (404) 873-8794
;             TELEFAX: (404) 873-8795
;   INFORMATION FOR SEQ ID NO: 5:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2032 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        FEATURE:
;             NAME/KEY:  misc. feature
;             LOCATION:  40..1926
;             OTHER INFORMATION:   /Function = "Nucleotides 40 through
;             1926 encode the amino acid sequence
;             for the Drosophila Melanogaster
;             Scavenger Receptor Class CI."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-241-581B-5

  Query Match             3.0%;  Score 51;  DB 4;  Length 2032;
  Best Local Similarity   57.4%;  Pred. No. 0.00032;
  Matches  112;  Conservative    0;  Mismatches   80;  Indels    3;  Gaps    1;

Qy       1358 CGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGA 1417
              | ||  || |  || | ||||    || | || || | | ||   |   || | ||   |
Db       1346 CAACAAAGGCCACAACTACAAAGCGAACAACAACCACTAAAAAACCGACAACAACTTCAA 1405

Qy       1418 CACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAA 1477
              || | ||   ||||||||||| ||    |  | |   |||  | ||  |||  | ||||
Db       1406 CAACGCCGAAGCCAACAACAACGAC---TTCAACCACACCAAAGTCTACAACTTCTACAA 1462

Qy       1478 CAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGA 1537
              |   |||   |  |  |  |||| ||||  ||| |||| ||   |||| |||||| ||
Db       1463 CGTCTACAACTTCAACAACACCAACGACAACAACTACAATAAATGTGTTTACAACAAAGA 1522

Qy       1538 CAATAACAGCAGAGA 1552
              || ||||  |  ||
Db       1523 AAACAACAATAATGA 1537


Search completed: June 19, 2003, 22:10:18
Job time : 93 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 21:25:52 ; Search time 285 Seconds
(without alignments)
8866.349 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	187.2	10.9	2556	9	US-09-938-842A-2418	Sequence 2418, Ap
	2	92.4	5.4	888	9	US-09-938-842A-453	Sequence 453, App
	3	55.8	3.2	2886	10	US-09-801-368-131	Sequence 131, App
c	4	54.8	3.2	4985	9	US-10-094-240-10	Sequence 10, Appl
	5	52	3.0	1236	9	US-10-077-584-3	Sequence 3, Appli
	6	51.6	3.0	867	10	US-09-216-393-340	Sequence 340, App
	7	51.6	3.0	867	10	US-09-216-393-342	Sequence 342, App
	8	51.6	3.0	1397	10	US-09-216-393-343	Sequence 343, App
c	9	51.6	3.0	1397	10	US-09-216-393-345	Sequence 345, App
	10	51.4	3.0	333	9	US-09-938-842A-3152	Sequence 3152, Ap
	11	45.8	2.7	2901	10	US-09-801-368-371	Sequence 371, App
	12	45.2	2.6	423	9	US-09-918-995-7147	Sequence 7147, Ap
	13	44.8	2.6	2277	10	US-09-801-368-223	Sequence 223, App
c	14	44	2.6	155074	9	US-10-026-188-6	Sequence 6, Appli
	15	43.4	2.5	1560	9	US-09-893-519A-82	Sequence 82, Appl
	16	43	2.5	1257	10	US-09-822-830A-428	Sequence 428, App
	17	43	2.5	2217	9	US-09-874-162A-6	Sequence 6, Appli
	18	43	2.5	2328	9	US-09-874-162A-9	Sequence 9, Appli
	19	43	2.5	2481	10	US-09-764-864-124	Sequence 124, App
	20	43	2.5	3150	9	US-09-938-842A-243	Sequence 243, App
	21	43	2.5	3545	10	US-09-764-864-144	Sequence 144, App
	22	43	2.5	4409	9	US-09-874-162A-7	Sequence 7, Appli
	23	43	2.5	4441	9	US-09-874-162A-4	Sequence 4, Appli
	24	43	2.5	4441	10	US-09-969-347-169	Sequence 169, App
c	25	43	2.5	6030	9	US-10-239-676-164	Sequence 164, App
c	26	42.6	2.5	456	10	US-09-864-761-4249	Sequence 4249, Ap
	27	42.2	2.5	58985	9	US-09-901-152-3	Sequence 3, Appli
	28	41.8	2.4	500	9	US-09-991-936-521	Sequence 521, App
	29	41.8	2.4	5085	9	US-10-198-846-9854	Sequence 9854, Ap
	30	41.4	2.4	493	10	US-09-864-761-5863	Sequence 5863, Ap
	31	41.2	2.4	14649	9	US-10-239-676-122	Sequence 122, App
	32	41	2.4	1530	10	US-09-799-777-113	Sequence 113, App
c	33	40.8	2.4	779	10	US-09-772-134B-69	Sequence 69, Appl
	34	40.8	2.4	2614	9	US-09-822-846-491	Sequence 491, App
c	35	40.8	2.4	6167	9	US-10-239-676-42	Sequence 42, Appl
	36	40.8	2.4	6604	10	US-09-880-107-1748	Sequence 1748, Ap
	37	40.8	2.4	335913	9	US-09-754-853A-2	Sequence 2, Appli
	38	40.8	2.4	335913	9	US-09-754-853A-3	Sequence 3, Appli
c	39	40.6	2.4	1075	10	US-09-864-761-19241	Sequence 19241, A
c	40	40.6	2.4	1403	10	US-09-864-761-2513	Sequence 2513, Ap
	41	40.4	2.3	2031	10	US-09-801-368-301	Sequence 301, App
	42	40.4	2.3	1830121	9	US-10-329-960-1	Sequence 1, Appli
	43	40.2	2.3	411	10	US-09-960-352-14521	Sequence 14521, A
c	44	40.2	2.3	17419	9	US-10-239-676-100	Sequence 100, App
c	45	40	2.3	6306	9	US-10-239-676-130	Sequence 130, App

ALIGNMENTS

RESULT 1
US-09-938-842A-2418
; Sequence 2418, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2418
; LENGTH: 2556
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2418

Query Match 10.9%; Score 187.2; DB 9; Length 2556;
Best Local Similarity 89.7%; Pred. No. 2.8e-37;
Matches 201; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
Qy        816 AATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAG 875
              | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2332 ACTGAAGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAG 2391

Qy        876 ATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTT 935
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2392 ATGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTT 2451

Qy        936 ATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGA 995
              ||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2452 ATATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGA 2511
```

```
Qy        996 TCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTA 1039
              ||||||||||||||     | | || | | |   |   ||  ||
Db       2512 TCAGAGGCTAGAAGGTTTGTTCATCATGACACCCCGTCATCATA 2555


RESULT 2
US-09-938-842A-453
; Sequence 453, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 453
;   LENGTH: 888
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-453

  Query Match             5.4%;  Score 92.4;  DB 9;  Length 888;
  Best Local Similarity   73.4%;  Pred. No. 2.4e-13;
  Matches  160;  Conservative    0;  Mismatches    1;  Indels   57;  Gaps    1;

Qy       1062 ATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTT 1121
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 ATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTT 60

Qy       1122 GCAGATTTTGAAGATCGCCAG---------------------------------------- 1142
              |||||||||||||||||||
Db         61 GCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAGTAGGCAA 120

Qy       1143 ------------------ATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAA 1184
                                |||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAAATGGTATACGATATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAA 180

Qy       1185 TTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAG 1222
              |||||||||||||||||||||||||||||||||||||
Db        181 TTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAG 218


RESULT 3
US-09-801-368-131
; Sequence 131, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 131
;   LENGTH: 2886
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-131

  Query Match             3.2%;  Score 55.8;  DB 10;  Length 2886;
  Best Local Similarity   54.0%;  Pred. No. 0.0011;
  Matches  114;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
              || || ||||||||||  || || || |  |||  |   |   || | |    ||   |
Db       1461 AGGCATCATCAACAACAACAGTAACAACGACGACGACGACAACAACAACAATAACGATAA 1520

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
              |   | | | ||||| || ||     ||  |   |   |    || |  ||||| |||||
Db       1521 CGATAACGATAACAATAATAGCAATAATAATAATAACAATAATAATAATAACAATAATAA 1580

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543
              | |    || ||   ||  | |  |||||||| ||||||  | | || ||    |||||
Db       1581 TAATAACAATAATAATAACAATAACAACAACAACAACAATAATAATAATAATAGTAATAA 1640

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAAT 1574
              |  | | | || ||   |||  |  | | |
Db       1641 TATTAAAAATAACGTTGATAACAATAATACT 1671


RESULT 4
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
;  APPLICANT: ZWIEBEL, LAURENCE J.
;  TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
;  FILE REFERENCE: N8289
;  CURRENT APPLICATION NUMBER: US/10/094,240
;  CURRENT FILING DATE:  2001-03-08
;  PRIOR APPLICATION NUMBER: 10/056,405
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: 60/264,649
;  PRIOR FILING DATE: 2001-01-26
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-094-240-10

  Query Match             3.2%;  Score 54.8;  DB 9;  Length 4985;
  Best Local Similarity   47.6%;  Pred. No. 0.0028;
  Matches  161;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy       1383 AATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACAAC 1442
              || || ||   |    ||     ||| |     ||| || ||  |  || |||||||||
Db       3063 AAAACAATAATAATAAACAATAATAAGAATAACAACAACAACAATAATAAGAACAACAAC 3004

Qy       1443 AGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCAAT 1502
              |      ||  |   | | |    || | ||| | |||||| |      || || ||||
Db       3003 AACAACAATAATAAGAATAATAATAATAACAATAGCAATAATAAGAATAATAATAACAAT 2944

Qy       1503 GACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATT 1562
                ||  || || || || |||  | | || ||  | ||||| |  |   | |||  ||
Db       2943 AGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAAAAAAAA 2884
```

```
Qy       1563 AAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTCTT 1622
Db       2883 TAATTATAATAAGAATAGTAATAATAATAATAATAATTAAAAATAATATTAATAATAA 2824

Qy       1623 TCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGATTA 1682
Db       2823 TAATAATAATAATAATAATAAGAATAATAATAATAATAATAATAATAATAATAATAATAA 2764

Qy       1683 CATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAA 1720
Db       2763 TAATAATAATCATAATAAAATTATTATTATGATAATAA 2726


RESULT 5
US-10-077-584-3
; Sequence 3, Application US/10077584
; Publication No. US20030073610A1
; GENERAL INFORMATION:
;  APPLICANT: LINDQUIST, SUSAN
;  APPLICANT:  KROBITSCH, SYLVIA
;  APPLICANT:  OUTEIRO, TIAGO F.
;  TITLE OF INVENTION: YEAST SCREENS FOR THE TREATMENT OF HUMAN DISEASE
;  FILE REFERENCE: ARCD:367US
;  CURRENT APPLICATION NUMBER: US/10/077,584
;  CURRENT FILING DATE:  2002-02-15
;  PRIOR APPLICATION NUMBER: 60/269,157
;  PRIOR FILING DATE: 2001-02-15
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1236
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(513)
US-10-077-584-3

  Query Match             3.0%;  Score 52;  DB 9;  Length 1236;
  Best Local Similarity   54.8%;  Pred. No. 0.0062;
  Matches  103;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1367 CACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCAC 1426
Db         58 CAGCAACAGCAACAACAGCAGCAACAGCAACAACAGCAGCAACAGCAACAACAGCAGCAA 117

Qy       1427 CACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACAT 1486
Db        118 CAGCAACAACAGCAGCAACAGCAACAACAGCAGCAACAGCAACAACAGCAGCAACAGCAA 177

Qy       1487 TGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAG 1546
Db        178 CAACAGCAGCAACAGCAACAACAGCAGCAACAGCAACAACAGCAGCAACAGCAACAACAG 237

Qy       1547 CAGAGACA 1554
Db        238 CAGCAACA 245


RESULT 6
US-09-216-393-340
; Sequence 340, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
;  APPLICANT: Milhausen, Michael James
;  TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: TX-1-C2
;  CURRENT APPLICATION NUMBER: US/09/216,393
;  CURRENT FILING DATE: 1998-12-18
;  EARLIER APPLICATION NUMBER: 08/994,825
;  EARLIER FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 340
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Toxoplasma gondii
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(867)
US-09-216-393-340

  Query Match             3.0%;  Score 51.6;  DB 10;  Length 867;
  Best Local Similarity   54.8%;  Pred. No. 0.0063;
  Matches  102;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy       1370 CATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCAC 1429
Db         485 CATCAACTACCACTACTACCACTACCACCACTACGACTACTACTACAACTACGACACCAA 544

Qy       1430 CAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGT 1489
Db         545 CAACAACTACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAACAACTA 604

Qy       1490 GGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAG 1549
Db         605 CACCAACAACAACGACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAA 664

Qy       1550 AGACAA 1555
Db         665 CGCCAA 670


RESULT 7
US-09-216-393-342
; Sequence 342, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
;  APPLICANT: Milhausen, Michael James
;  TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: TX-1-C2
;  CURRENT APPLICATION NUMBER: US/09/216,393
;  CURRENT FILING DATE: 1998-12-18
;  EARLIER APPLICATION NUMBER: 08/994,825
;  EARLIER FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 342
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Toxoplasma gondii
US-09-216-393-342

  Query Match             3.0%;  Score 51.6;  DB 10;  Length 867;
  Best Local Similarity   54.8%;  Pred. No. 0.0063;
  Matches  102;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy       1370 CATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCAC 1429
Db         485 CATCAACTACCACTACTACCACTACCACCACTACGACTACTACTACAACTACGACACCAA 544

Qy       1430 CAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGT 1489
Db         545 CAACAACTACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAACAACTA 604

Qy       1490 GGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAG 1549
Db         605 CACCAACAACAACGACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAA 664

Qy       1550 AGACAA 1555
Db         665 CGCCAA 670
```

RESULT 8
US-09-216-393-343
; Sequence 343, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 343
; LENGTH: 1397
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)..(1104)
US-09-216-393-343

Query Match 3.0%; Score 51.6; DB 10; Length 1397;
Best Local Similarity 54.8%; Pred. No. 0.0085;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 1370 CATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCAC 1429
Db 722 CATCAACTACCACTACTACCACTACCACCACTACGACTACTACTACAACTACGACACCAA 781
Qy 1430 CAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGT 1489
Db 782 CAACAACTACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAACAACTA 841
Qy 1490 GGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAG 1549
Db 842 CACCAACAACAACGACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAA 901
Qy 1550 AGACAA 1555
Db 902 CGCCAA 907

RESULT 9
US-09-216-393-345/c
; Sequence 345, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 345
; LENGTH: 1397
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
US-09-216-393-345

Query Match 3.0%; Score 51.6; DB 10; Length 1397;
Best Local Similarity 54.8%; Pred. No. 0.0085;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 1370 CATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCAC 1429
Db 676 CATCAACTACCACTACTACCACTACCACCACTACGACTACTACTACAACTACGACACCAA 617
Qy 1430 CAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGT 1489
Db 616 CAACAACTACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAACAACTA 557
Qy 1490 GGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAG 1549
Db 556 CACCAACAACAACGACAACAACCACAACAACTACACCAACAACAACGACAACAACCACAA 497
Qy 1550 AGACAA 1555
Db 496 CGCCAA 491

RESULT 10
US-09-938-842A-3152
; Sequence 3152, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3152
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3152

Query Match 3.0%; Score 51.4; DB 9; Length 333;
Best Local Similarity 90.2%; Pred. No. 0.0039;
Matches 55; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1003 CTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAA 1062
Db 80 CTTCCAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGA 139
Qy 1063 T 1063
Db 140 T 140

RESULT 11
US-09-801-368-371
; Sequence 371, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric

```
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 371
;   LENGTH: 2901
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-371

  Query Match             2.7%;  Score 45.8;  DB 10;  Length 2901;
  Best Local Similarity   59.7%;  Pred. No. 0.41;
  Matches   77;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy       1427 CACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACAT 1486
              ||   |||   ||||       | |  | | |||  ||   |||| |||||||||| | ||
Db       1633 CAAGCACAGGCACAAGCACAAGCACAGGCGCAGGCACAACAACAACAACAACAACAGCAA 1692

Qy       1487 TGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAG 1546
                |  |   |||  |||  |||||||| ||||     |||||||  |||| |||||
Db       1693 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAG 1752

Qy       1547 CAGAGACAA 1555
              |||   |||
Db       1753 CAGCAGCAA 1761


RESULT 12
US-09-918-995-7147
; Sequence 7147, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7147
;   LENGTH: 423
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(423)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-7147

  Query Match             2.6%;  Score 45.2;  DB 9;  Length 423;
  Best Local Similarity   49.7%;  Pred. No. 0.18;
  Matches   86;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy       1466 AAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGT 1525
              ||| |  || || || || |      |  |    ||  | | ||| || || || ||
Db        215 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274

Qy       1526 TGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTT 1585
                | || ||  | || || |  | | | |||  ||  ||| |  |||         | |||
Db        275 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAANNNNNNNNCCTTT 334

Qy       1586 ATGATACCGATTTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTT 1638
              |  |     ||| |             ||| || ||| |||||||| | |||
Db        335 TTTTTTTTTTTTTTNNNNNNNNNNANAATTTTTTTTTTTTAAAAAAATTTTT 387


RESULT 13
US-09-801-368-223
; Sequence 223, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 223
;   LENGTH: 2277
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-223

  Query Match             2.6%;  Score 44.8;  DB 10;  Length 2277;
  Best Local Similarity   52.7%;  Pred. No. 0.63;
  Matches   97;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              ||| ||    | | |  |  | |  || |   |  || ||  || |  || || |||
Db        805 CAATCACCAGCACAGCCCCAGCAATCATCTCAACAGCAGATCCAGCAGCCTCAGCATCAA 864

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              |  || || ||       |  |   |||  ||   |||| ||||||||| ||||
Db        865 CCCCAGCATCAACCGCAACAGCAACAGCAACAGCAACAACAACAACAACAGCAACAG 924

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 1552
                |   |||   ||  | || ||| | |||     ||| |||  |||  ||||||||  |
Db        925 CAACAGCAACAGCAACAGCAGCAACAGCAACAGCAACAACATCAACAGCAGCAA 984

Qy       1553 CAAA 1556
              || |
Db        985 CAGA 988


RESULT 14
US-10-026-188-6/c
; Sequence 6, Application US/10026188
; Patent No. US20020164645A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
;  TITLE OF INVENTION:  Ion Channel
;  FILE REFERENCE: 02307E-114910US
;  CURRENT APPLICATION NUMBER: US/10/026,188
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: US 60/259,379
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 8
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 155074
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human genomic region containing ltrpc5
;   OTHER INFORMATION: (Human Chromosome 11p15.5 PAC clone pDJ915f1
;   OTHER INFORMATION: containing KvLQT1 gene)
US-10-026-188-6

  Query Match              2.6%;  Score 44;  DB 9;  Length 155074;
  Best Local Similarity   59.7%;  Pred. No. 14;
  Matches   74;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy       1366 CCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCA 1425
             ||||||||| || |  ||| |||| |  |   ||    |   | | | | |||||  |
Db      45767 CCACCATCACCACCACCAAAACCACCACCACCAACACCATCACCATCACCGCCACCAAAA 45708

Qy       1426 CCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACA 1485
              |||||| ||||| |||||  |          |||    |  || ||||| || ||  ||||
Db      45707 CCACCACCAACACCAACACCGCCTCCATCACCACCACCAAAACCACCACCACCATCAACA 45648

Qy       1486 TTGT 1489
              |||
Db      45647 CTGT 45644


RESULT 15
US-09-893-519A-82
; Sequence 82, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
;  APPLICANT: ANADYS PHARMACEUTICALS, INC.
;  APPLICANT:  THOMPSON, Craig
;  APPLICANT:  MOORE, Jeffrey
;  APPLICANT:  BUURMAN, Ed T.
;  APPLICANT:  BRADLEY, John
;  APPLICANT:  DESILVA, Thamara
;  APPLICANT:  HARRIS, Sandra
;  APPLICANT:  KOMARNITSKY, Svetlana
;  APPLICANT:  MENDILLO, Marc
;  APPLICANT:  MOORE, Daniel
;  APPLICANT:  MCCOY, Melissa
;  APPLICANT:  SANDERSON, Karen
;  APPLICANT:  HAQ, Tariq
;  APPLICANT:  ZHU, Shuhao
;  APPLICANT:  LONG, Fan
;  APPLICANT:  DAVIDOV, Eugene
;  TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
;  FILE REFERENCE: 0342/1G548-US2
;  CURRENT APPLICATION NUMBER: US/09/893,519A
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: US 60/215,164
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 60/224,457
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 146
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
;   LENGTH: 1560
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: Genbank/CAA96830.1
;   DATABASE ENTRY DATE: 1997-08-11
;   RELEVANT RESIDUES: (1)..(1560)
US-09-893-519A-82

  Query Match              2.5%;  Score 43.4;  DB 9;  Length 1560;
  Best Local Similarity   49.8%;  Pred. No. 1.1;
  Matches  110;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy       1330 TGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCA 1389
              ||   |||   ||| || || | |      ||     || || || ||||  ||| | ||
Db        285 TGCGCAAAGCTTGGGACAATCGGATATCGCACAACAGCAACAACAGCAACAACAACAACA 344

Qy       1390 TCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGG 1449
              |  | |||  |  || || |||     || || || |  ||||| |||| ||
Db        345 GCAACCAGACATCGCGCAACAGCAACCTCAACAGCAACCTCAACAGCAACCTCAACAGCA 404

Qy       1450 ATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAA 1509
              | |    |||  ||   |||| |||| ||| | ||   |  |      |   ||| |
Db        405 ACCTCAACAGCAACCTCAACAGCAACCTCAACAGCAACCTCAACAGCAACCTCAACAGCA 464

Qy       1510 ACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550
              ||  |||   |||      || | |  || |      ||||
Db        465 ACCTCAACTTCAACCACTTCAGCCACAACTAGGGACCCAGA 505

Search completed: June 19, 2003, 23:44:22
Job time : 291 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:03:46 ; Search time 4023 Seconds
(without alignments)
10761.960 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1722	100.0	1722	33	US-09-890-220-1	Sequence 1, Appli
2	1722	100.0	1722	33	US-09-890-475-57	Sequence 57, Appl
3	1720.4	99.9	1722	33	US-09-890-220-47	Sequence 47, Appl
4	1569.6	91.1	1715	33	US-09-890-220-4	Sequence 4, Appli
5	1569.6	91.1	1715	33	US-09-890-475-58	Sequence 58, Appl
6	1537.6	89.3	1737	33	US-09-890-220-7	Sequence 7, Appli
7	979.8	56.9	1497	19	US-09-513-996A-69038	Sequence 69038, A
8	412.2	23.9	6338	33	US-09-890-220-3	Sequence 3, Appli
9	366.8	21.3	1722	42	US-10-219-999-5373	Sequence 5373, Ap
10	326.4	19.0	606	22	US-09-565-309A-32201	Sequence 32201, A
11	319.4	18.5	5895	33	US-09-890-220-6	Sequence 6, Appli
12	319.4	18.5	201471	20	US-09-534-859-587	Sequence 587, App
13	319.4	18.5	201471	31	US-09-803-736-587	Sequence 587, App
14	314.2	18.2	319	19	US-09-513-996A-69042	Sequence 69042, A
15	314	18.2	2346	74	US-60-303-460-9	Sequence 9, Appli
16	306.8	17.8	638	22	US-09-565-309A-52491	Sequence 52491, A
17	306.8	17.8	640	22	US-09-565-309A-32200	Sequence 32200, A
18	306.8	17.8	640	22	US-09-565-309A-61982	Sequence 61982, A
19	306.8	17.8	640	22	US-09-595-329A-945	Sequence 945, App
20	306.4	17.8	640	19	US-09-513-996A-41806	Sequence 41806, A
21	291.4	16.9	2280	17	US-09-339-947A-2	Sequence 2, Appli

	22	284.2	16.5	2249	74	US-60-303-460-1	Sequence 1, Appli
	23	264.8	15.4	2145	74	US-60-303-460-13	Sequence 13, Appl
	24	257.4	14.9	1841	42	US-10-219-999-16339	Sequence 16339, A
	25	257.4	14.9	1841	75	US-60-312-544-4687	Sequence 4687, Ap
	26	257.4	14.9	1856	33	US-09-865-439A-20106	Sequence 20106, A
	27	257.4	14.9	2459	74	US-60-303-460-5	Sequence 5, Appli
	28	255	14.8	2276	74	US-60-303-460-7	Sequence 7, Appli
	29	252.4	14.7	2248	17	US-09-339-947A-9	Sequence 9, Appli
	30	247.8	14.4	593	22	US-09-565-309A-32199	Sequence 32199, A
	31	243.6	14.1	2296	74	US-60-303-460-11	Sequence 11, Appl
	32	242.8	14.1	677	25	US-09-654-617-26623	Sequence 26623, A
	33	242.8	14.1	677	27	US-09-684-016-26623	Sequence 26623, A
	34	236.2	13.7	2640	17	US-09-398-237-6	Sequence 6, Appli
	35	234	13.6	303	25	US-09-654-617-163164	Sequence 163164,
	36	234	13.6	303	27	US-09-684-016-163164	Sequence 163164,
	37	230.8	13.4	558	33	US-09-890-220-13	Sequence 13, Appl
	38	221.8	12.9	704	74	US-60-303-460-17	Sequence 17, Appl
	39	214	12.4	1176	74	US-60-303-460-3	Sequence 3, Appli
	40	210.8	12.2	600	43	US-10-260-238-2870	Sequence 2870, Ap
	41	210.8	12.2	600	61	US-60-170-912-179	Sequence 179, App
	42	210.8	12.2	600	76	US-60-325-448-2870	Sequence 2870, Ap
c	43	209.8	12.2	990	25	US-09-654-617-386601	Sequence 386601,
c	44	209.8	12.2	990	27	US-09-684-016-386601	Sequence 386601,
	45	209	12.1	531	33	US-09-874-708A-29223	Sequence 29223, A

ALIGNMENTS

```
RESULT 1
US-09-890-220-1
; Sequence 1, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-1

  Query Match           100.0%;  Score 1722;  DB 33;  Length 1722;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
             |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 2
US-09-890-475-57
; Sequence 57, Application US/09890475
; GENERAL INFORMATION:
;  APPLICANT: Johanson, Urban
;  APPLICANT:  West, Joanne
;  APPLICANT:  Dean, Caroline
;  TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,475
;  CURRENT FILING DATE:  2001-11-13
;  PRIOR APPLICATION NUMBER: PCT/GB00/00197
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: GB 9902660.1
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Landsberg
;   OTHER INFORMATION: erecta VRN2 cDNA
US-09-890-475-57

  Query Match             100.0%;  Score 1722;  DB 33;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
```

```
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380

Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
Db       1681 TACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 3
US-09-890-220-47
; Sequence 47, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-47

  Query Match             99.9%;  Score 1720.4;  DB 33;  Length 1722;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1721;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTTT 60

Qy         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120
Db         61 CAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGTT 120

Qy        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180
Db        121 GCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTTG 180

Qy        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240
Db        181 CACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGGC 240

Qy        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300
Db        241 AGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTGA 300

Qy        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360
Db        301 TATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCAT 360

Qy        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420
Db        361 CGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGAT 420

Qy        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480
Db        421 CTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAG 480

Qy        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540
Db        481 TTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGC 540

Qy        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600
Db        541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAAT 600

Qy        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660
Db        601 ACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTG 660

Qy        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720
Db        661 ATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAG 720

Qy        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780
Db        721 GTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTT 780

Qy        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840
Db        781 TAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGAT 840

Qy        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900
Db        841 ATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCA 900

Qy        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960
Db        901 TAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCC 960

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
Db        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
Db       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
Db       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
Db       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGAA 1200

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
Db       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320
Db       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGA 1320

Qy       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
Db       1321 GATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACT 1380
```

```
Qy       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAACA 1440

Qy       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCCCA 1500

Qy       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAA 1560

Qy       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCTTC 1620

Qy       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACTGAT 1680

Qy       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722
             |||||||||||||||||||||||||||||||||||||||||||
Db       1681 TACATTTTTAGTTAAAAAAAAAAAATGGATTCTGCTTATAACT 1722


RESULT 4
US-09-890-220-4
; Sequence 4, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1715
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-4

  Query Match             91.1%;  Score 1569.6;  DB 33;  Length 1715;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   10;  Gaps    4;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              ||||||||||||||||||||||  |||||||||||||| ||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||||| ||||| |||||||||||| |||| ||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718
              |||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720

Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838
              |||| |||||||||||||||||||||  |||| || ||||||||||||||||||||||||
Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900

Qy        899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958
              |||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy        959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy       1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy       1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy       1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy       1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy       1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318
              || ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy       1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380

Qy       1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
```

```
Db       1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy       1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498

Db       1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy       1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558

Db       1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy       1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618

Db       1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy       1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677

Db       1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAT 1705

Db       1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701


RESULT 5
US-09-890-475-58
; Sequence 58, Application US/09890475
; GENERAL INFORMATION:
;  APPLICANT: Johanson, Urban
;  APPLICANT:  West, Joanne
;  APPLICANT:  Dean, Caroline
;  TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,475
;  CURRENT FILING DATE:  2001-11-13
;  PRIOR APPLICATION NUMBER: PCT/GB00/00197
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: GB 9902660.1
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
;   LENGTH: 1715
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
;   OTHER INFORMATION: cDNA
US-09-890-475-58

  Query Match             91.1%;  Score 1569.6;  DB 33;  Length 1715;
  Best Local Similarity   96.5%;  Pred. No. 0;
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   10;  Gaps    4;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58

Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118

Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178

Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238

Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298

Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358

Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418

Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478

Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 538

Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCT 540

Qy        539 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGA 598

Db        541 GCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGA 600

Qy        599 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 658

Db        601 ATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAG 660

Qy        659 TGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAG 718

Db        661 TGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGTAGACAAAG 720

Qy        719 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 778

Db        721 AGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAG 780

Qy        779 TTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGG 838

Db        781 TTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGTGGTTTAGG 840

Qy        839 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGC 898

Db        841 ATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGC 900

Qy        899 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGT 958

Db        901 CATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGT 960

Qy        959 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1018

Db        961 CCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACT 1020

Qy       1019 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1078

Db       1021 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGT 1080

Qy       1079 AATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1138

Db       1081 AATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCG 1140

Qy       1139 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1198

Db       1141 CCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTG 1200

Qy       1199 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1258

Db       1201 GAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGA 1260

Qy       1259 AGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTG 1318

Db       1261 AGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTGTTG 1320

Qy       1319 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1378

Db       1321 GAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAA 1380
```

```
Qy       1379 CTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
             ||||||||||||||||||||||||||||||||| |||||       ||| | |||||||||
Db       1381 CTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACTAACAACAACAA 1434

Qy       1439 CAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1498
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATCATCC 1494

Qy       1499 CAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAAGT 1558
              |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||
Db       1495 GAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAC-AAGT 1553

Qy       1559 AATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTATTCT 1618
             |||||||||||||  |||||| ||  |||||||||| ||||||||||||||||||||||
Db       1554 AATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTGTAACTTATTCT 1613

Qy       1619 TCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCAACT 1677
             |||||||| |||||||||||||||||||||||  |||||| ||||||||||||||||||
Db       1614 TCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAGTGTATTCAACT 1673

Qy       1678 GATTACATTTTTAGTTAAAAAAAAAAAT 1705
             |||||||||||||||||||||||   ||
Db       1674 GATTACATTTTTAGTTAAAAAAATGGAT 1701


RESULT 6
US-09-890-220-7
; Sequence 7, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 1737
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-7

  Query Match             89.3%;  Score 1537.6;  DB 33;  Length 1737;
  Best Local Similarity   95.3%;  Pred. No. 0;
  Matches 1649;  Conservative    0;  Mismatches   49;  Indels   32;  Gaps    5;

Qy          1 CAAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCT 58
              ||||||||||||||||||||||  |||||||||||| |||||||||||||||||||||||
Db          1 CAAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCT 60

Qy         59 TTCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 118
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTG 120

Qy        119 TTGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 178
              |||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTT 180

Qy        179 TGCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAG 238
              |||||||||||||||||| ||||| |||||||||| |||| ||||||||||||||||||
Db        181 TGCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAG 240

Qy        239 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTT 300

Qy        299 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCC 360

Qy        359 ATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAG 418
              ||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        361 ATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAG 420

Qy        419 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGA 478
              |||||||||||||||||||||||||||||||||||||||| |||||||| | ||||||
Db        421 ATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGA 480

Qy        479 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCA------- 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGTGGG 540

Qy        532 --------------AGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 576
                           ||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 600

Qy        577 TTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 636
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 660

Qy        637 TCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 696
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        661 TCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGA 720

Qy        697 AACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 756
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCT 780

Qy        757 TTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTA 816
              |||||||||||||||||||||||| ||||||||||||||||||  |||| || |
Db        781 TTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGA 840

Qy        817 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGA 900

Qy        877 TGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 936
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTA 960

Qy        937 TATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 996
              |||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGAT 1020

Qy        997 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCC 1080

Qy       1057 AGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACG 1116
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1081 AGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACG 1140

Qy       1117 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATG 1200

Qy       1177 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATG 1260

Qy       1237 GTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTT 1296
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db       1261 GTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTT 1320

Qy       1297 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTG 1380

Qy       1357 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAG 1416
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
```

```
Db       1381 TCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG 1440

Qy       1417 ACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACA 1476
                 ||| | ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       1441 ------TCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACA 1494

Qy       1477 ACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGG 1536
              |||||||||||||||||||||| |||||||||| |||||||||||||||||||||||||
Db       1495 ACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGG 1554

Qy       1537 ACAATAACAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGAT 1596
              |||||||||||||||| |||||||||||||||||  |||||| ||  ||||||||||
Db       1555 ACAATAACAGCAGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGAT 1613

Qy       1597 TTATCGGATTGTAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTT 1655
               |||||||||||||||||||||||||||| |||||||||||||||||||||||||  ||
Db       1614 CTATCGGATTGTAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATT 1673

Qy       1656 TTATATGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAT 1705
              |||| |||||||||||||||||||||||||||||||||||||||   ||
Db       1674 TTATTTGTTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 1723


RESULT 7
US-09-513-996A-69038
; Sequence 69038, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69038
;   LENGTH: 1497
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..1497
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..1497 / Ceres Seq. ID 2185972
US-09-513-996A-69038

  Query Match             56.9%;  Score 979.8;  DB 19;  Length 1497;
  Best Local Similarity   89.7%;  Pred. No. 1.1e-218;
  Matches 1118;  Conservative    0;  Mismatches   27;  Indels  102;  Gaps    2;

Qy        355 ACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGT 414
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGT 379

Qy        415 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAA 439

Qy        475 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 CTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGG 499

Qy        535 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCG 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db        500 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGTACATCT 559

Qy        595 AAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAG 654
                | | | |  ||  | |
Db        560 CTGTAAATCCAACCATAA---------------------------------------- 577

Qy        655 GAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGAC 714
                                    | | |  | ||||||||||||||||||||||||
Db        578 --------------------------TGTTTTGCAGCTCGAAACCTCGTAAGCGGAGAC 610

Qy        715 AAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 AAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCAC 670

Qy        775 CCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 CCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTT 730

Qy        835 TAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCAC 894
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        731 TAGGATATCCCGAGGCAACAGAACTTGCTGGACAATTTGAGATGACCAGCAACATTCCAC 790

Qy        895 CAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTG 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 CAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTG 850

Qy        955 TGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGA------- 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 TGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGTACTG 910

Qy       1008 --------------------------AGCCACCTACTTCTTCAGAAACGCCAATTCTATC 1041
                                         |||||||||||||||||||||||||||||||||
Db        911 CCATATTAACCGAGAAAATTTCTTCCAGCCACCTACTTCTTCAGAAACGCCAATTCTATC 970

Qy       1042 ATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGG 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 ATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGG 1030

Qy       1102 ATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGG 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 ATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGG 1090

Qy       1162 ATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAA 1221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 ATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAA 1150

Qy       1222 GGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGA 1281
              |||||||||||||||||||| || ||||||||||||||||||||||||||||||||||||
Db       1151 GGGTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGAGA 1210

Qy       1282 AAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTAT 1341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 AAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTAT 1270

Qy       1342 GGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATT 1401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 GGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATT 1330

Qy       1402 GCCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTG 1461
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 GGCGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTG 1390

Qy       1462 ACTCAAACACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACA 1521
              |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1391 ACTCAAACACCAACAACAATAACATTGTGAGTCATCCCAATGACATAAACAACAAGAACA 1450

Qy       1522 ATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAG 1568
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       1451 ATGTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAATAG 1497


RESULT 8
US-09-890-220-3
; Sequence 3, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics us
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
```

```
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 6338
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-3

  Query Match             23.9%;  Score 412.2;  DB 33;  Length 6338;
  Best Local Similarity   98.8%;  Pred. No. 2.1e-85;
  Matches  414;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1304 ACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363
             | | | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5102 AATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 5161

Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5162 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 5221

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5222 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 5281

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543
             ||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5282 CATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 5341

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5342 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 5401

Qy       1604 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5402 ATTGTAACTTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGT 5461

Qy       1664 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5462 TAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATAACT 5520


RESULT 9
US-10-219-999-5373
; Sequence 5373, Application US/10219999
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)C
;  CURRENT APPLICATION NUMBER: US/10/219,999
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 60/324,109
;  PRIOR FILING DATE: 2001-09-21
;  PRIOR APPLICATION NUMBER: US 60/312,544
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 5373
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (164)..(1450)
;   OTHER INFORMATION:
US-10-219-999-5373

  Query Match             21.3%;  Score 366.8;  DB 42;  Length 1722;
  Best Local Similarity   59.8%;  Pred. No. 5.7e-75;
  Matches  757;  Conservative    0;  Mismatches  422;  Indels   87;  Gaps    5;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
              || |||||  |||| |||| ||  |   |      |    || ||||  ||| || |||||
Db        161 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 220

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
              ||||  || ||||| ||||| || ||||||  ||| || || ||  |  |||  ||  ||
Db        221 GAGAGCCTTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCT 280

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
              ||    || || || ||||||   ||||| |||   || |||||  ||||||  || |||
Db        281 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAA 340

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
              || | ||  |||    | || || ||  |||||| |||| |||  |  | ||||   |
Db        341 AGGAGGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTT 397

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
              |  || |||||||| |  || || | |||||||||| |||||||  ||||  ||||| | |
Db        398 CGAAAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAAC 457

Qy        528 TTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG 587
              || |||||  ||| |||||||  |   |||||| |||||| |||  | |||||||||  |
Db        458 TTTAAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGG 517

Qy        588 CTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG 647
              ||   ||||| ||||| |||| ||||| |||||||||| || ||   || | ||  |||
Db        518 GTTACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGATATATTGAGATCAGAG 577

Qy        648 GAAGAAGGAAGTGATGACGAT---AAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGT 704
              | |  |    ||  |    |    |||   |  ||||||   |||| || | ||||||
Db        578 AATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCTCGA 637

Qy        705 AAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCG 764
              || || |||  ||  |          | |      |   |    || |||   ||  |   |
Db        638 AAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTGGAG 697

Qy        765 TTGGATTCACCCA---------------------------------------------- 777
              ||||| |||||
Db        698 TTGGACTCACCAGAAGGCATACATAATGGATTTCTACAAAAAGATGATGATATCTTATCC 757

Qy        778 ----------------GTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGA 820
                              || ||   ||  |  ||  |   || |  ||||| |  | ||
Db        758 TGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTCCCTAGTGGAAGA 817

Qy        821 T--------GGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTT 872
                     |||||   ||||   |  ||||  |  | |  ||       |  ||  |
Db        818 AATGACGGAGGAAAATTTGGTCCTGATCATCCTGGCACCATGGACAACCTGGAGCATGTG 877

Qy        873 GAGATGACCAGCAACATTCCACCA---------GCCATAGCCCACTCTTCTCTGGACGCT 923
              ||    |   ||||||||||           |||||  |||| |||||| ||||| ||
Db        878 GAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCTTCTGTGGACCCT 937

Qy        924 GGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTA 983
              |    ||       | |    |    | |  |  |   || |||| |||| ||| ||
Db        938 GAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCTAAAACAAAGAAGCTA 997

Qy        984 TCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCAT 1043
                || ||||||||  | |||| ||   | ||||| ||||| | |  |||| |||
Db        998 AGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAGAGACTTTTCTTTCAC 1057

Qy       1044 TCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGAT 1103
              || |||||||||||||| ||||| || || |||||  | || ||||| ||||| || ||
Db       1058 TCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGACCGTGATAGTGAAGAC 1117

Qy       1104 GAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGAT 1163
              ||||| ||||||||  ||||||||  ||||||||  |   |||||||||  || |||||||||
Db       1118 GAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTTGACGATTTTGTGGAT 1177
```

```
Qy       1164 GTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGG 1223
              ||     |||||||||||| ||  |||||||||| |||||||| ||| | || || ||||||
Db       1178 GTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGG 1237

Qy       1224 GTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAA 1283
              ||  |||||||||||||| |   | ||||| ||||| |||||||| |.  || | |  |||
Db       1238 GTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAA 1297

Qy       1284 GAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1343
              ||| ||  |   |   || |  | || ||||||||||| || ||  |||| ||||| |||
Db       1298 GAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTCATGATCAAACTTTGG 1357

Qy       1344 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGC 1403
              || ||||| ||| | ||  |     || || |||||||| |  |    |  | || | || |
Db       1358 AATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATAGTATTAGATAGTTAC 1417

Qy       1404 CGTAAT 1409
                | |||
Db       1418 AGAAAT 1423
```

RESULT 10
US-09-565-309A-32201
; Sequence 32201, Application US/09565309A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-0853P
; CURRENT APPLICATION NUMBER: US/09/565,309A
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 68449
; SEQ ID NO 32201
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(606)
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: (1)..(606)
; OTHER INFORMATION: 39533:110036 (Clone Number:Unique Sequence Identifier)
US-09-565-309A-32201

Query Match 19.0%; Score 326.4; DB 22; Length 606;
Best Local Similarity 98.0%; Pred. No. 1.1e-65;
Matches 341; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy       1373 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAA 318

Qy       1433 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGA 378

Qy       1493 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAG- 1551
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGA 438

Qy       1552 ACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAAC 1611
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        439 ACAAAGTAATTAAATAGGAAAATCTCCGGCTTAGATGATACCGATTTATCGGATTGTAAC 498

Qy       1612 TTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTAT 1671
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTATTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTAT 558

Qy       1672 TCAACTGATTACATTTTTAGTTAAAAAAAAAAATGGATTCTGCTTATA 1719
              |||||||||||||||||||||||||||||||| |||||||||  | |
Db        559 TCAACTGATTACATTTTTAGTTAAAAAAAAAAAAGGATTCTGCAAAAA 606
```

RESULT 11
US-09-890-220-6
; Sequence 6, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: Gendall, Anthony
; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
; TITLE OF INVENTION: vernalisation gene VRN2.
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5895
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-890-220-6

Query Match 18.5%; Score 319.4; DB 33; Length 5895;
Best Local Similarity 92.7%; Pred. No. 1.2e-63;
Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

```
Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4677 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 4736

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||||||| |||||       |||
Db       4737 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 4790

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       4791 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 4850

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||
Db       4851 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 4910

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||| |||||||||||||||||  |||||| ||  |||||||||| ||||||||||
Db       4911 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 4969

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAG 1666
              |||||||||||||||||||  |||||||||||||||||||||||||  |||||| ||||||
Db       4970 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 5029

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||   ||
Db       5030 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 5068
```

RESULT 12
US-09-534-859-587
; Sequence 587, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534,859
; CURRENT FILING DATE: 2000-03-29

```
;  NUMBER OF SEQ ID NOS:       1127
; SEQ ID NO 587
;   LENGTH: 201471
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-534-859-587

  Query Match            18.5%;  Score 319.4;  DB 20;  Length 201471;
  Best Local Similarity   92.7%;  Pred. No. 4.6e-63;
  Matches  370;  Conservative    0;  Mismatches   21;  Indels    8;  Gaps    3;

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42342 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 42401

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||||||||| |||||       |||
Db      42402 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 42455

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      42456 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 42515

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db      42516 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 42575

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||| |||||||||||||||||  |||||| ||  ||||||||||| ||||||||||
Db      42576 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 42634

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTTATATGTTAG 1666
              |||||||||||||||||| |||||||||||||||||||||  |||||| ||||||
Db      42635 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 42694

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||||    ||
Db      42695 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 42733


RESULT 13
US-09-803-736-587
; Sequence 587, Application US/09803736
; GENERAL INFORMATION:
;  APPLICANT: Bush, David F.
;  APPLICANT:  Levin, Irena M.
;  APPLICANT:  Norris, Susan R.
;  APPLICANT:  Rounsley, Steven D.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
;  FILE REFERENCE: 38-10(15493)D
;  CURRENT APPLICATION NUMBER: US/09/803,736
;  CURRENT FILING DATE:  2001-03-12
;  PRIOR APPLICATION NUMBER: US 09/534,859
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: identified by Attorney Docket number 04983.0206CPUS01 38-10
;  PRIOR FILING DATE: 2000-10-20
;  NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 587
;   LENGTH: 201471
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-803-736-587

  Query Match            18.5%;  Score 319.4;  DB 31;  Length 201471;
  Best Local Similarity   92.7%;  Pred. No. 4.6e-63;
  Matches  370;  Conservative    0;  Mismatches   21;  Indels    8;  Gaps    3;

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      42342 TCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 42401

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACC 1427
              |||||||||||||||||||||||||||||||||||||||||||||| |||||       |||
Db      42402 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------TCACT 42455

Qy       1428 ACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATT 1487
              | |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      42456 AACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATT 42515

Qy       1488 GTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGC 1547
              |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||
Db      42516 GTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGC 42575

Qy       1548 AGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTG 1607
              |||||| |||||||||||||||||  |||||| ||  ||||||||||| ||||||||||
Db      42576 AGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGGATTG 42634

Qy       1608 TAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTTATATGTTAG 1666
              |||||||||||||||||| |||||||||||||||||||||  |||||| ||||||
Db      42635 TAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTGTTAG 42694

Qy       1667 TGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||||    ||
Db      42695 TGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 42733


RESULT 14
US-09-513-996A-69042
; Sequence 69042, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69042
;   LENGTH: 319
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..319
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..319 / Ceres Seq. ID 2185983
US-09-513-996A-69042

  Query Match            18.2%;  Score 314.2;  DB 19;  Length 319;
  Best Local Similarity   99.1%;  Pred. No. 6.3e-63;
  Matches  316;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1375 ACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACA 1434
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1 ACAACTGCAATACCATCCTCGAGAATTGGCGTAATAGCTCAGACACCACCACCACCAACA 60

Qy       1435 ACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGATC 1494
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db         61 ACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGAGTC 120

Qy       1495 ATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACA 180

Qy       1555 AAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTA 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGGATTGTAACTTA 240

Qy       1615 TTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCA 1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTTCTTTCTTAAAAAATTGTTTAGGAGCAAACAAATTTTTTATATGTTAGTGTATTCA 300

Qy       1675 ACTGATTACATTTTTAGTT 1693
              |||||||||||||||||||
```

```
Db        301 ACTGATTACATTTTTAGTT 319


RESULT 15
US-60-303-460-9
; Sequence 9, Application US/60303460
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT:  Butler, Karlene
;  APPLICANT:  Niu, Xiping
;  TITLE OF INVENTION: Flowering Regulating Polynucleotides and Polypeptides
;  FILE REFERENCE: BB1493 US PRV
;  CURRENT APPLICATION NUMBER: US/60/303,460
;  CURRENT FILING DATE:  2001-07-06
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
;   LENGTH: 2346
;   TYPE: DNA
;   ORGANISM: Helianthus sp.
US-60-303-460-9

  Query Match             18.2%;  Score 314;  DB 74;  Length 2346;
  Best Local Similarity   58.4%;  Pred. No. 1.5e-62;
  Matches  598;  Conservative    0;  Mismatches  405;  Indels   21;  Gaps    2;

Qy        384 TACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAAC 443
              |||||   ||   | |  | |     ||  |     ||||  |  || ||||| ||
Db        944 TACAACGATGCTCCTAGACCCGGTGTAATGAGATTGCGGTCTGGAAATGTCGTTTTTAAT 1003

Qy        444 TATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCA 503
              || ||| | |  || || ||||| || |  ||||||||||  ||||||| |||||||||
Db       1004 TACAAGTACTACAACAATACATTGCATAGGACTGAAGTTACAGAGGATTTTTCTTGTCCG 1063

Qy        504 TTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCATCTCAT 563
              || || |   ||  ||| | ||| | |||||| |||  |||||| |||| |||||| |||
Db       1064 TTCTGTTTGGTGAAATGCGCTAGTTACAAGGGCCTGAGATTTCACTTGACTTCATCACAT 1123

Qy        564 GATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTA 623
              ||| | ||| | | |||||||   | ||   |||||  |  |||   || | ||| ||| |
Db       1124 GATCTCTTTCATTATGAGTTTTGGGTTACTGAAGACTGTCAGGTCGTGATTGTATCTATC 1183

Qy        624 AAACTTAATTCCTTCATATTTGAGGAAG---AAGGAAGTGATGACGATAAATTTGAGCCC 680
              | |    || |   ||  | ||||  |       |  |  | |||    |||    |
Db       1184 AGAACCGATGCAAGCAGCTCTGAGATAATTCCTGAGAACGTTGATCCGAAACAACAAATG 1243

Qy        681 TTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGG 740
              ||||||   |||   ||| |  | |  ||  |  ||  |         || | | | |
Db       1244 TTCTCTTATTGCCATAAATCGAGAAGACGCCGGAAACCAAAAACTCAAACTCAAAACGCA 1303

Qy        741 AGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAAT 800
              |  | |      .          || || ||| |    |   |     ||     |
Db       1304 AATCAT-----------------GTGCATCCACTCGTGCTGGATTCAACCATGTCTACA 1345

Qy        801 GGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTT 860
              |   |||     ||||  |||    |    |||| |||   |  |  ||     ||
Db       1346 GCTCTCAGTGAGCTTATAGATAATACAGATGGTGTAGCTGAGTGCATGGAGCTGGATGCA 1405

Qy        861 GCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGAC 920
              | | |    ||  |   |||    |    ||   |||  | || |  ||   | |
Db       1406 TGTAGTCCGGATGCAAGTGCCACATGCCCCTCAAATGCCGAACCCGAGCCTATACAGTCA 1465

Qy        921 GCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAG 980
                ||  |     |   |  ||      |||  || | |    | | ||||| ||||||
Db       1466 TTACCTGAAAGCAGCAACCTTGCACCTCCAGCCATGCTTCAGTTTGCAAAGACGAGAAAG 1525

Qy        981 TTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTAT 1040
              ||||| | |||||||||| ||  |||||| ||      || || |||||||| || |||| |
Db       1526 TTATCGGTTGAGCGATCCGACCCTAGAAACCGAGCCCTGCTGCAGAAACGGCAGTTCTTT 1585

Qy       1041 CATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAG 1100
              || || ||| |||  |||||||||||| | || |||||| |  |||| || || || ||
Db       1586 CACTCCCACCGAGCTCAGCCAATGGCGATAGAACAAGTATTCGCTGAGCGAGACAGTGAA 1645

Qy       1101 GATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTG 1160
              |||||||| || || |||||||| ||| | || || ||   |||||| || ||||| |||
Db       1646 GATGAAGTTGACGATGATGTTGCTGATCTCGAGGACCGTAGGATGCTGGACGACTTCGTG 1705

Qy       1161 GATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAA 1220
              |||||    ||||| || || ||  | |||||||| |||||||| ||||| ||||| |||
Db       1706 GATGTTTCCAAAGACGAGAAACAGATGATGCATCTATGGAACTCATTTGTGAGAAAGCAA 1765

Qy       1221 AGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAG 1280
              ||||  | || ||||| ||||| || ||| | || ||||| |||||||    |  | |
Db       1766 CGGGTGCTGGCGGATGGCCATATTTCGTGGTCGTGCGAAGCGTTTTCAAAGCTGCATGGT 1825

Qy       1281 AAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTA 1340
              ||||  |    |      ||||||| |||||||| |||||||||||| |||| ||| ||
Db       1826 CAAGACCTTGTTCAGGCTCCATCACTACTCTGGTGCTGGAGATTGTTTATGATAAAATTA 1885

Qy       1341 TGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAAT 1400
              ||||| ||||| ||  | |||||    | | | ||||| || |||   ||||| ||  |
Db       1886 TGGAATCATGGGCTCCTAGACTCGCGGAGCTTAAACAATTGTAATGTGATCCTTGAACAA 1945

Qy       1401 TGCC 1404
              | ||
Db       1946 TACC 1949


Search completed: June 19, 2003, 23:17:45
Job time : 4034 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:25:25 ; Search time 2561 Seconds
(without alignments)
10889.742 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	628.2	36.5	672	10	AV822548	AV822548 AV822548
c	2	266.6	15.5	526	17	CNS00WYB	AL094193 Arabidops
	3	237.8	13.8	680	10	AW038171	AW038171 EST279828
c	4	232.4	13.5	769	14	BQ505017	BQ505017 EST612432
	5	226	13.1	1128	11	AY104964	AY104964 Zea mays
	6	212	12.3	846	12	BG648271	BG648271 EST509890
	7	206.2	12.0	480	9	AJ469021	AJ469021 AJ469021
	8	203	11.8	569	14	BQ611550	BQ611550 sap61f09.
	9	201.4	11.7	600	13	BI479743	BI479743 WHE3451_E
	10	198.4	11.5	474	10	AW234600	AW234600 sf17b08.y
c	11	197.8	11.5	391	10	AV817525	AV817525 AV817525
	12	190	11.0	561	14	BQ740672	BQ740672 saq50e05.
	13	188.8	11.0	467	12	BG580495	BG580495 EST482220
	14	176.4	10.2	364	9	AI164598	AI164598 A065P57U
	15	174.6	10.1	405	12	BG314044	BG314044 WHE2467_D
	16	173.6	10.1	554	13	BI784785	BI784785 saf93c12.
	17	171.2	9.9	546	13	BI321750	BI321750 saf98e10.
	18	170.2	9.9	444	12	BG551103	BG551103 sad33c03.
	19	169.6	9.8	612	10	BE203328	BE203328 EST403350
	20	164	9.5	429	14	BQ459233	BQ459233 HA01O05r
	21	160.6	9.3	445	10	AW202144	AW202144 sf12d08.y
	22	154.6	9.0	326	9	AI163743	AI163743 A047P36U
c	23	144.2	8.4	561	14	BQ162202	BQ162202 WHE0751_H
c	24	143	8.3	662	14	BQ279630	BQ279630 1091030A1
c	25	132.8	7.7	519	14	BQ168835	BQ168835 WHE2467_D
	26	126	7.3	644	13	BJ453006	BJ453006 BJ453006
c	27	125.6	7.3	322	10	AW309408	AW309408 sf17b08.x
c	28	124	7.2	757	13	BJ460553	BJ460553 BJ460553
	29	121	7.0	515	12	BF481866	BF481866 FM1_22_D0
c	30	120.6	7.0	236	14	BQ904049	BQ904049 Ta03_18d1
	31	113.6	6.6	560	9	AI795561	AI795561 614009G07
	32	113.6	6.6	590	12	BG649311	BG649311 EM1_77_A0
	33	113.6	6.6	594	10	AW000259	AW000259 614009G07
c	34	110.6	6.4	586	14	BQ243123	BQ243123 TaE15019A
	35	109.2	6.3	727	14	BQ801572	BQ801572 WHE2816_A
	36	108.6	6.3	732	14	BQ841916	BQ841916 WHE2986_G
c	37	105.8	6.1	142	10	AV783345	AV783345 AV783345
	38	104.8	6.1	463	14	C72616	C72616 C72616 Rice
	39	101.2	5.9	576	10	AW076157	AW076157 614063H03
	40	100.4	5.8	675	13	BI932726	BI932726 EST552615
c	41	94.8	5.5	523	13	BI992698	BI992698 1020066F1
c	42	94	5.5	553	12	BG909756	BG909756 TaLr1108b
c	43	94	5.5	553	14	BQ620168	BQ620168 TaLr1108B
	44	94	5.5	593	13	BM377912	BM377912 EBem04_SQ
	45	93.6	5.4	540	10	BE587348	BE587348 WHE661_B0

ALIGNMENTS

RESULT 1
AV822548
LOCUS AV822548 672 bp mRNA linear EST 01-APR-2002
DEFINITION AV822548 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-08-K19 5', mRNA sequence.
ACCESSION AV822548
VERSION AV822548.1 GI:19864597
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 672)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998).This clone is in a

modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .672
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL05-08-K19"
                     /clone_lib="RAFL5"
                     /dev_stage="rosette plants"
                     /lab_host="SOLR"
                     /note="Site_1: SstI; Site_2: XhoI; subjected to
                     dehydration-treated(1,2,5,10,24 hr)"
BASE COUNT      197 a    116 c    143 g    216 t
ORIGIN

  Query Match             36.5%;  Score 628.2;  DB 10;  Length 672;
  Best Local Similarity   97.0%;  Pred. No. 3.8e-129;
  Matches  651;  Conservative    0;  Mismatches   18;  Indels    2;  Gaps    1;

Qy          2 AAGCTTCTTCAATTTTGCTTG--CTCTCTCTTACACAGCCAATCGGTGTTTTCGCAGCTT 59
              |||||||||||||||||||||  ||||||||||||| |||||||||||||||||||||||
Db          2 AAGCTTCTTCAATTTTGCTTGCTCTCTCTCTTACACGGCCAATCGGTGTTTTCGCAGCTT 61

Qy         60 TCAGGCCTCAATCCAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGT 119
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TCAGGCCTCAATACAAGACATTCTATATAAGCATATTGCAGAAGAGGCGGTTCTAATTGT 121

Qy        120 TGCATTGAGTTTATCGCTATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTT 179
              ||||| ||||| | |  ||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGCATGGAGTTGAACAATATGACGTAGGGAAATTCTAATTTAGGGGAGGCCTCAGAGTTT 181

Qy        180 GCACTAACTTCATAATCGGCTCTTGACGTTGTTGAGTGTAATTGAACAAGAATGTGTAGG 239
              |||||||||||||||| ||||| |||||||||| |||| ||||||||||||||||||||
Db        182 GCACTAACTTCATAATCAGCTCTGGACGTTGTTGATTGTATTTGAACAAGAATGTGTAGG 241

Qy        240 CAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAGAATCTCTTG 301

Qy        300 ATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 ATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTAGGCAACCCA 361

Qy        360 TCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCAAGA 419
              |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        362 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTCAAGA 421

Qy        420 TCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAA 479
              |||||||||||||||||||||||||||||||||||||||| |||||||| | |||||||
Db        422 TCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAA 481

Qy        480 GTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTG 541

Qy        540 CAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAA 599
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        542 CAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTGGAAGAA 601

Qy        600 TACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGT 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGT 661

Qy        660 GATGACGATAA 670
              ||||| |||||
Db        662 GATGATGATAA 672


RESULT 2
CNS00WYB/c
LOCUS       CNS00WYB                526 bp    DNA     linear   GSS 28-JUN-1999
DEFINITION  Arabidopsis thaliana genome survey sequence T7 end of BAC T12M8 of
```

TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.

```
ACCESSION   AL094193
VERSION     AL094193.1  GI:5295347
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 526)
  AUTHORS   Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P.,
            Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 526)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="T12M8"
                     /clone_lib="TAMU"
                     /note="end : T7"
BASE COUNT      148 a     71 c     90 g    217 t
ORIGIN

  Query Match             15.5%;  Score 266.6;  DB 17;  Length 526;
  Best Local Similarity   92.1%;  Pred. No. 8.1e-49;
  Matches  316;  Conservative    0;  Mismatches   19;  Indels    8;  Gaps    3;

Qy       1364 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCAC 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        526 AGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCAG------T 473

Qy       1424 CACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAA 1483
              ||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        472 CACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAA 413

Qy       1484 CATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAA 1543
              |||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||
Db        412 CATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAA 353

Qy       1544 CAGCAGAGACAAAGTAATTAAATAGGAAAATCTCCGGCTTTTATGATACCGATTTATCGG 1603
              |||||||||| |||||||||||||||  |||||| ||  |||||||||| ||||||
Db        352 CAGCAGAGAC-AAGTAATTAAATAGGAAACACTCCGGTTTAGATGATACCGATCTATCGG 294

Qy       1604 ATTGTAACTTATTCTTCTTTCTT-AAAAAATTGTTTAGGAGCAAACAAATTTTTTATATG 1662
              |||||||||||||||||||||||| |||||||||||||||||||||||  |||||| ||
Db        293 ATTGTAACTTATTCTTCTTTCTTAAAAAAATTGTTTAGGAGCAAACAAAGATTTTATTTG 234

Qy       1663 TTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAAAAAAT 1705
              |||||||||||||||||||||||||||||||||||||||||   ||
Db        233 TTAGTGTATTCAACTGATTACATTTTTAGTTAAAAAAATGGAT 191


RESULT 3
AW038171
LOCUS       AW038171                680 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST279828 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
            clone cLET1N5, mRNA sequence.
ACCESSION   AW038171
VERSION     AW038171.1  GI:5896925
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
```

```
            Lycopersicon.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
            Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
            Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni
            ,J.
  TITLE     Generation of ESTs from tomato leaf tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Lycopersicon esculentum"
                     /cultivar="Rio Grande PtoR"
                     /db_xref="taxon:4081"
                     /clone="cLET1N5"
                     /clone_lib="tomato mixed elicitor, BTI"
                     /tissue_type="leaf"
                     /dev_stage="4-6 week old plants"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLET - Inoculated with a variety of disease response
                     elicitors. Plants  exposed to 2,6 dichloroisonicotinic
                     acid, BTH, jasmonic acid, ethylene,  fenthion, EIX,
                     okadaic acid, or systemin prior to tissue harvest.  EcoRI
                     site was destroyed during cloning."
BASE COUNT      192 a    120 c    164 g    204 t
ORIGIN

  Query Match           13.8%;  Score 237.8;  DB 10;  Length 680;
  Best Local Similarity   74.1%;  Pred. No. 2.1e-42;
  Matches  315;  Conservative    0;  Mismatches  107;  Indels    3;  Gaps    1;

Qy        987 GCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCT 1046
              | |||||| || ||  | ||||  |    ||| || || ||| | ||||||| |||||||
Db          1 GTTGAGCGCTCTGATCCCAGAAATCGTGCACTCCTGCAAAAAAGGCAATTCTTTCATTCT 60

Qy       1047 CACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAA 1106
              || || | |||||| ||||| || |||||||| | || ||||| ||'|| |||||||||
Db         61 CATAGGGCCCAGCCCATGGCACTGGAGCAAGTTTTATCGGACCGAGACAGTGAGGATGAA 120

Qy       1107 GTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTG 1166
              || ||||| |||||||||||| ||||||||||   |||||||||||| ||||||||||||
Db         121 GTTGATGATGATGTTGCAGATCTTGAAGATCGAAGGATGCTTGATGATTTTGTGGATGTG 180

Qy       1167 AATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTT 1226
              |  |||||||||||||| | |||||||| |||||||| ||||| ||||| ||||||||
Db         181 ACCAAAGATGAAAAGCAAGTGATGCATCTGTGGAACTCATTTGTTAGAAAGCAAAGGGTG 240

Qy       1227 ATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAG 1286
              | |||||||||||||| |||||||||||||| || |||||||   |  | |   |  |
Db         241 TTGGCAGATGGTCATATCCCTTGGGCATGTGAGGCCTTTTCAAAGCTGCATGGTCAGAGG 300

Qy       1287 TTGCACCGTTACTCA---TCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1343
              ||   ||       ||    |||| | | |||||||||||||| ||  |||| || || |||
Db         301 TTTGCCCAAGCACCAGCCACACTATGCAGGTGTTGGAGATTATTCATGATGAAGCTGTGG 360

Qy       1344 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGC 1403
              |||||||| ||||| ||  |    || || ||||| || ||    ||  | ||| | | |
Db         361 AACCATGGCCTTGTTGATGCGCGTACAATTAACAATTGTAACCTAATATTAGAGCAGTTC 420

Qy       1404 CGTAA 1408
              |  ||
Db         421 CAAAA 425


RESULT 4
BQ505017/c
LOCUS       BQ505017                769 bp    mRNA    linear   EST 22-JUL-2002
DEFINITION  EST612432 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
            3' end, mRNA sequence.
ACCESSION   BQ505017
VERSION     BQ505017.2  GI:21920982
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21363886.
            Other_ESTs: EST612431
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMGB40"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating eyes
                     , tubers, or roots."
BASE COUNT      239 a    176 c    140 g    214 t
ORIGIN

  Query Match           13.5%;  Score 232.4;  DB 14;  Length 769;
  Best Local Similarity   73.0%;  Pred. No. 3.4e-41;
  Matches  327;  Conservative    0;  Mismatches  116;  Indels    5;  Gaps    2;

Qy        961 CTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTC 1020
              | | | | ||| |||||||||||||| | |||||| ||     | ||||  |    ||| |
Db        747 CAGTTGCCAAGTCAAGAAAGTTATCAGTTGAGCGCTCTATCCCCAGAAATCGTGCACTCC 688

Qy       1021 TTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAA 1080
              | || ||| | ||| ||| ||||||||| || | |||||| ||||| || ||||||||
Db        687 TGCAAAAAAGGCAA-TCTTTCATTCTCATAGGGCCCAGCCCATGGCACTGGAGCAAGTTT 629

Qy       1081 TGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCC 1140
              | || || || || || || |||||||||| ||||| |||||||||||| |||||||||
Db        628 TATCAGATCGAGACAGTGAGGATGAAGTTGATGATGATGTTGCAGATCTTGAAGATCGAA 569

Qy       1141 AGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGA 1200
              |||||||||||| |||||||||||| |||||||||||||| | |||||||| ||||
Db        568 GGATGCTTGATGATTTTGTGGATGTGACCAAAGATGAAAAGCAAATGATGCATCTGTGGA 509

Qy       1201 ACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAG 1260
              |||| ||||| ||||| |||||||| | |||||||||||| ||| ||||||||||||| |
Db        508 ACTCATTTGTTAGAAAGCAAAGGGTGTTGGCAGATGGTCACATCCCTTGGGCATGTGAGG 449

Qy       1261 CATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTAC----TCATCACTCTTCTGGTGT 1316
              | |||||||   |  | ||  |||  |   ||    |     | ||  | | | |||||
Db        448 CCTTTTCAAAGCTGCATGATTCAGATGTTTGCCCAAGCACCAGCCTCGTTGTGCAGGTGT 389
```

```
Qy       1317 TGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAAC 1376
              |||||||| ||  |||| ||  | ||||||||||| ||||| ||  |   ||  || |||
Db        388 TGGAGATTATTCATGATGAAGTTGTGGAACCATGGCCTTGTTGATGCGCGTACAATTAAC 329

Qy       1377 AACTGCAATACCATCCTCGAGAATTGCC 1404
              || || ||    ||  | ||| | | ||
Db        328 AATTGTAACCTAATATTAGAGCAGTTCC 301


RESULT 5
AY104964
LOCUS       AY104964               1128 bp    mRNA    linear   HTC 25-MAY-2002
DEFINITION  Zea mays PCO105783 mRNA sequence.
ACCESSION   AY104964
VERSION     AY104964.1  GI:21208042
KEYWORDS    HTC.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 1128)
  AUTHORS   Coe,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .1128
                     /organism="Zea mays"
                     /db_xref="MaizeDB:636128"
                     /db_xref="taxon:4577"
                     /clone="PCO105783"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      355 a    199 c    242 g    332 t
ORIGIN

  Query Match           13.1%;  Score 226;  DB 11;  Length 1128;
  Best Local Similarity   57.3%;  Pred. No. 9.4e-40;
  Matches  527;  Conservative    0;  Mismatches  330;  Indels   63;  Gaps    4;

Qy        485 GGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAATT 544
              ||| ||||  |||||||||||||||||| | | | |||||| |||||||||||  ||   ||
Db         11 GGAAGATTTCTCTTGTCCATTTTGCTATGTACGATGTGGAAGCTTCAAGGGTCTAGGATG 70

Qy        545 TCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACCA 604
              ||||| || ||||| |||||| ||||  | | ||||||   | | |  ||||| |||||
Db         71 CCATTTAAACTCATCACATGATCTATTCCACTATGAGTTTTGGATATCTGAAGAGTACCA 130

Qy        605 GACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG---GAAGAAGGAAGTGA 661
              |   |||||||||  | | ||   | || | |  | |   |||      |  |   | |
Db        131 GGTTGTTAATGTTAGTCTGAAGGCTGATGCTTGGAGAACAGAGCTTTTTGCGGAGGGCGT 190

Qy        662 TGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGG 721
              |||    |  | |  | || |||    |||| |   |  |||||  |||| | |
Db        191 TGATCCAAGGCATCAAACATTTTCTTATCGCTCAAGGTTTAAGAAGCGTAGACGATCAAA 250

Qy        722 TGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTT 781
                || ||   | |  |  || |   | |||  ||         |||| |||
Db        251 GAACACAATGGAGAAAATCAGGC---ATGTACACTCACATATTATGGAATCAGGTTCACC 307
```

```
Qy        782 AACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGATA 841
              |||   ||| |  ||   || | |  |  | |  |  ||  |   ||  | |
Db        308 TGAAGATGAGGCAGGATCTGAGGACAACTTTGTGCAAGGGGAGAATGGGACTTCTGTAGC 367

Qy        842 TCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCAT 901
                |  | |  || | ||||  |   ||  |   | ||        ||||||  |  |
Db        368 AAATGCTTCGATTGATCCTGCTCAATCTTTACATGGCAGCAATCTTTCACCACCAACAGT 427

Qy        902 AGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCC 961
              |  || | |
Db        428 ACTACAGTTTGGG---------------------------------------------- 440

Qy        962 TGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCT 1021
                  |||||||| ||| |||   ||||| |||| ||  ||||||  |  | ||| ||
Db        441 -------AAGACAAGGAAGCTAT---CTGAGAGATCTGACCCTAGAAATCGGCAACTCCT 490

Qy       1022 TCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAAT 1081
              || ||||| || ||||  |||||||||| |  |||||||||   || ||||||||  |
Db        491 GCAAAAACGACAGTTCTTCCATTCTCACAGGGCGCAGCCAATGCAACTGGAGCAAGTGTT 550

Qy       1082 GTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCA 1141
              || ||||| ||||| || |||||||| ||||| ||| |||| || || || ||| |
Db        551 CTCGGACCGTGATAGTGAAGATGAAGTTGATGATGATATTGCTGACTTCGAGGATAGAAG 610

Qy       1142 GATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAA 1201
              |||||||||||| |||| |||| |  |||||||||| |   | |||||| | ||||
Db        611 AATGCTTGATGATTTTGTTGATGTTACGAAAGATGAAAAACTTATTATGCATATGTGGAA 670

Qy       1202 CTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGC 1261
              || |||||  |||||||||| ||  |||| |||||||||||  ||||||| || || ||
Db        671 TTCATTTGTTCGAAAACAAAGAGTGTTAGCTGATGGTCATATACCTTGGGCCTGCGAGGC 730

Qy       1262 ATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAG 1321
              ||| ||    ||  | |   | |  |    |  | |  | || | ||| ||||| |
Db        731 ATTCTCCCAGTTGCATGGACGACAACTTATACAAAATCCTGCTCTGCTGTGGGGTTGGCG 790

Qy       1322 ATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTG 1381
              || ||  |||||||||| |||||||||    || | ||  |   ||| || ||||  ||
Db        791 TTTCTTCATGATTAAACTTTGGAACCATAACATTTTAGATGCCCGCACTATGAACACATG 850

Qy       1382 CAATACCATCCTCGAGAATT 1401
              ||||||  ||||  | | ||
Db        851 CAATACAGTCCTTCAAATTT 870


RESULT 6
BG648271
LOCUS       BG648271               846 bp    mRNA    linear   EST 24-APR-2001
DEFINITION  EST509890 HOGA Medicago truncatula cDNA clone pHOGA-19E20 5' end,
            mRNA sequence.
ACCESSION   BG648271
VERSION     BG648271.1  GI:13783383
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
            Utterback,T., Cho,J. and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula treated with
            oligogalacturonides of DP 6-20
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Michael G. Hahn
            Complex Carbohydrate Research Center
            University of Georgia
            220 Riverbend Road, Athens, GA 30602-4712, USA
            Tel: 706-542-4457
            Fax: 706-542-4412
            Email: hahn@ccrc.uga.edu
            G391877e TIGR sequence name: MTMCJ34TK More information is
```

```
          available at: www.medicago.org
          Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Medicago truncatula"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pHOGA-19E20"
                     /clone_lib="HOGA"
                     /tissue_type="3 day old seedling roots"
                     /dev_stage="24 hours after treatment in the dark at 26 C
                     with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
                     presence of 100 ug/ml Gentamicin"
                     /lab_host="XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in SOLR cells."
BASE COUNT      248 a    161 c    185 g    252 t
ORIGIN

  Query Match             12.3%;  Score 212;  DB 12;  Length 846;
  Best Local Similarity   64.2%;  Pred. No. 1.2e-36;
  Matches  352;  Conservative    0;  Mismatches  190;  Indels    6;  Gaps    2;

Qy        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290
              |||||  |||| ||||  |  | | |     ||    || ||||  ||| || |||||||
Db         52 ATGTGCCGGCAAAATTCACCGGTGCATAATTCCGGTGAAGAAGAAATTACAGCTGATGAA 111

Qy        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350
              | ||| ||||| ||||| |||||||||  | | || || ||  |  | | ||| ||||
Db        112 AGTCTTTTGATTTATTGCAAACCTGTTGAATTCTACAATATTCTCTATCGTCGTTCTCTT 171

Qy        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410
                ||| ||  | ||||||   ||||| |||  ||| |||||   ||||||||  |||||
Db        172 CACAATCCTACTTTTCTTAAGAGATGTTTGCGCTATAAAATAAAAGCAAAGCAAAAAAGG 231

Qy        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470
              | ||   ||    | || || || ||||||||||||| ||| ||| | ||  |  | |
Db        232 AGGTTGCGA---GCAGGAATTGTGGTTTTCAACTATAGGGACTGTTACAATGCGCTTCGA 288

Qy        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530
              || |||||||| |  || || | ||||||||||||||||||  ||||  ||||  |||||
Db        289 AAGACTGAAGTGACTGAAGACTTTTCTTGTCCATTTTGCTTGATGCAGTGTGCGAGCTTT 348

Qy        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590
              |||||  ||| |||||||  |    | |||| |||||| ||||  | |||||||||  | |
Db        349 AAGGGTTTGCGATTTCATCTTTGTGCATCACATGATCTATTCAACTTTGAGTTCTGGGTC 408

Qy        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650
                | | || |||||  |||| || || |||||||||  | |  .  || | ||  ||| |
Db        409 ACCAATGATTACCAAGCAGTGAACGTCTCTGTAAAAACTGACATATTGAGATCAGAGAAT 468

Qy        651 GAAGGAAGTGATGACGAT---AAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAG 707
              |  |    ||  |    |    |||   | ||||||   |||| || | | |||| |||
Db        469 GTTGCTGATGGAGTAATTCCACAATCACAAACCTTCTTCTTCTGTTCAAGATCTCGAAAG 528

Qy        708 CGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTG 767
              || |||| ||||||    |    | ||    | |     || |||   ||  |   ||||
Db        529 CGTAGACCAAGAGGCTCCGTTCAAAATGGGAAGCATGCCAATGTAAATTTCCTGAAGTTG 588

Qy        768 GATTCACC 775
              ||||||||
Db        589 GATTCACC 596


RESULT 7
AJ469021
LOCUS       AJ469021                 480 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AJ469021 S00008 Hordeum vulgare cDNA clone S0000800053C04F1, mRNA
            sequence.
ACCESSION   AJ469021
VERSION     AJ469021.1  GI:21184977
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
  TITLE     Barley EST's
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Schulman AH
            Institute of Biotechnology
            University of Helsinki
            P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
            Finland.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Hordeum vulgare"
                     /db_xref="taxon:4513"
                     /clone="S0000800053C04F1"
                     /clone_lib="S00008"
                     /tissue_type="Callus"
                     /note="Callus K19"
BASE COUNT      138 a     95 c    111 g    136 t
ORIGIN

  Query Match             12.0%;  Score 206.2;  DB 9;  Length 480;
  Best Local Similarity   68.9%;  Pred. No. 2.1e-35;
  Matches  283;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        999 GAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAG 1058
              ||  | |||| ||  .| ||| || |||||||| || || |  ||||||||||| |  |||
Db          1 GATCCCAGAAACCGACAACTCCTGCAGAAACGTCAGTTTTTCCATTCTCACAGGGCACAG 60

Qy       1059 CCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGAT 1118
              |||||||| || || |||||  | || || || ||||| || |||||||| ||||| ||
Db         61 CCAATGGCACTGGAACAAGTTTTCTCGGATCGTGATAGTGAAGATGAAGTTGATGATGAC 120

Qy       1119 GTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAA 1178
                | |||||||||||||| | | |||||||||||| ||||| ||||| |  || |||||
Db        121 ATCGCAGATTTTGAAGATAGACGGATGCTTGATGATTTTGTTGATGTCACGAATGATGAG 180

Qy       1179 AAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGT 1238
              || |   | |||||| | |||| || ||||  | ||||||||||  |||| ||||||
Db        181 AAACTTATTATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTGCTAGCTGATGGT 240

Qy       1239 CATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTAC 1298
              |||||  ||||||| |||| |||| || |  || | |  ||| |  |    |  |
Db        241 CATATTCCTTGGGCCTGTGAGGCATTCTCCCGGCTTCATGGAAAACATCTTGTACAGAAT 300

Qy       1299 TCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTC 1358
                |  | ||  | ||| | ||| | ||  || |||||||||| ||||||||  | ||  |
Db        301 CCTCCTCTACTATGGAGCTGGCGTTTCCTTATGATTAAACTGTGGAACCACAGTCTATTA 360

Qy       1359 GACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAAT 1409
              ||  |     |||| ||   |||   || || ||  |    | ||  |||
Db        361 GATGCCCGTGCCATGAATGTCTGCGGCACAATTCTTCAAGGCTACCAAAAT 411


RESULT 8
BQ611550
LOCUS       BQ611550                 569 bp    mRNA    linear   EST 26-JUN-2002
DEFINITION  sap61f09.y1 Gm-c1087 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1087-5250 5' similar to TR:Q9ZNT9 Q9ZNT9
            FERTILIZATION-INDEPENDENT SEED 2 PROTEIN. ;, mRNA sequence.
ACCESSION   BQ611550
VERSION     BQ611550.1  GI:21601219
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 569)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 440.
FEATURES Location/Qualifiers
source 1. .569
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1087-5250"
/clone_lib="Gm-c1087"
/tissue_type="Roots"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The mRNA was prepared using polyatract mRNA system from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site (5'GAGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTTT). EcoRI adapters (5'OH-AATTCGGCACGAG and 3'GCCGTGCTCp) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."
BASE COUNT 172 a 106 c 122 g 169 t
ORIGIN

Query Match 11.8%; Score 203; DB 14; Length 569;
Best Local Similarity 69.6%; Pred. No. 1.1e-34;
Matches 275; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
Qy        964 CTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTC 1023
              || |||| ||||  ||| ||      || ||||||||  | |||| ||   | ||||| |
Db         50 CTGCTAAAACAAAGAAGCTAAGCATGGATCGATCAGACTCAAGAAACCGGATGCTTCTGC 109

Qy       1024 AGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGT 1083
              ||||  | |  |||| ||| || ||||||||||||||| ||||| || || |||||  | |
Db        110 AGAAGAGACTTTTCTTTCACTCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTAT 169

Qy       1084 CTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGA 1143
              | ||||| ||||| || || ||||| ||||||||  |||||||| |||||||| |   ||
Db        170 CAGACCGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGA 229

Qy       1144 TGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACT 1203
              |||||||||| ||||||||||     |||||||||| ||  |||||||||| |||||||
Db        230 TGCTTGATGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACT 289

Qy       1204 CGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCAT 1263
```

```
              | ||| | || || ||||||||  | ||||||||||| |  | ||||| ||||| ||||
Db        290 CTTTTATGAGGAAGCAAAGGGTGCTGGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCAT 349

Qy       1264 TTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGAT 1323
              |||| |   || | |  |||||| ||  |   |   |  |  | || ||||  |
Db        350 TTTCCAAGCTTCATGGAAAAGAGCTGATCTCATCTCCTGCTTTATTTTGGTACTCCTATC 409

Qy       1324 TGTTTTTGATTAAACTATGGAACCATGGACTTGTC 1358
              |    |||||     | |     |||||| || ||
Db        410 TTCCATTGATATTTATCTAATTTCATGGAGTTTTC 444
```

RESULT 9
BI479743
LOCUS BI479743 600 bp mRNA linear EST 28-AUG-2001
DEFINITION WHE3451_E10_I19ZS Wheat pre-anthesis spike cDNA library Triticum aestivum cDNA clone WHE3451_E10_I19, mRNA sequence.
ACCESSION BI479743
VERSION BI479743.1 GI:15323794
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 600)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES Location/Qualifiers
source 1. .600
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE3451_E10_I19"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
BASE COUNT 165 a 126 c 143 g 166 t
ORIGIN

Query Match 11.7%; Score 201.4; DB 13; Length 600;
Best Local Similarity 70.0%; Pred. No. 2.5e-34;
Matches 271; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
Qy       1004 TAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAAT 1063
              | | ||||  | |||||| |||||||| || ||||  |||||||||| |  ||||||||
Db        214 TGGCAGCCGACAACTTCTGCAGAAACGTCAGTTCTTCCATTCTCACAGGGCACAGCCAAT 273
```

```
Qy       1064 GGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGC 1123
Db        274 GGCAGTGGAACAAGTTTTGTCAGATCATGACAGCGAGGATGAAGTTGACGATGACATTGC 333

Qy       1124 AGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCA 1183
Db        334 CGACTTGGAAGATAGACGGATGCTTGCTGATTTTCTTGATGTCACGAAAGATGAGAAGCT 393

Qy       1184 ATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATAT 1243
Db        394 TATTATGCATATGTGGAATTCATTTATTCGGAAACAAAGGGTGCTAGCTGATGGGCATAT 453

Qy       1244 CTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATC 1303
Db        454 ACCTTGGGCCTGTGAGGGCTTCTCCCGGCTTCATGGACCGCAGCTTGTACAAAACCCTCC 513

Qy       1304 ACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTC 1363
Db        514 TCTGCTCTGGTGCTGGCGTTTGGTGATGATTAAGCTGTGGAACCACAGCCTGCTGGATGC 573

Qy       1364 AGCCACCATCAACAACTGCAATACCAT 1390
Db        574 GCGCGCCATGAACACATGCAACACCAT 600
```

RESULT 10
AW234600

```
LOCUS       AW234600                474 bp    mRNA    linear   EST 17-JUL-2000
DEFINITION  sf17b08.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1028-208 5' similar to TR:O23525 O23525 HYPOTHETICAL 27.2 KD
            PROTEIN. ;, mRNA sequence.
ACCESSION   AW234600
VERSION     AW234600.1  GI:6566973
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1799   Std Error: 0.00
            High quality sequence stop: 412.
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-208"
                     /clone_lib="Gm-c1028"
                     /tissue_type="roots of 'Supernod' plants"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; The mRNA was isolated from roots of Glycine max
                     'Supernod' plants generously donated by Dr. Gary Stacey.
                     The seedlings were innoculated with Bradyrhizobium
                     japonicus, strain USDA110 priot to harvest. Stratagene's
                     cDNA synthesis Kit (catalog number 200401) was used to
                     synthesize the cDNA.  First-strand synthesis was performed
                     with 5-methyl dCTP, hence the ligated cDNA was
                     hemimethylated.  A modification of Stratagene's
                     first-strand synthesis primer was used.  An 'anchor'
                     nucleotide (V=A,C, or G) was added to the 3' end of the
                     primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor
                     the primer at the 5' end of the poly(A) tract.  After
                     second-strand synthesis, the cDNA ends were filled in with
                     cloned Pfu DNA polymerase, ligated to EcoRI adapters and
                     subsequently phosphorylated.  The XhoI site within the
                     first-strand synthesis primer was then restricted by
                     digestion with XhoI; all XhoI sites in the cDNA would be
                     protected by their hemimethylated status.  The cDNA
                     constructs were size-fractionated with a 500bp cutoff,
                     using GibcoBRL Life Technologies' cDNA Size Fractionation
                     column.  The column eluent was then ligated into
                     Stratagene's pBluescript II XR Predigested vector
                     (pBluescript II SK(+) that has been digested with EcoRI
                     and XhoI, and phosporylated by Stratagene).  Both the
                     white and blue colonies appear to contain recombinant
                     plasmids with cDNA inserts, based on size (n=25).  This
                     library was constructed by Dr. Paul Keim and  Dr. Virginia
                     Coryell."
BASE COUNT      133 a     86 c    107 g    148 t
ORIGIN
```

```
  Query Match           11.5%;  Score 198.4;  DB 10;  Length 474;
  Best Local Similarity  68.4%;  Pred. No. 1.1e-33;
  Matches  290;  Conservative    0;  Mismatches  131;  Indels    3;  Gaps    1;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
Db         40 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 99

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
Db        100 GAGAGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCT 159

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
Db        160 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAA 219

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
Db        220 AGGAGGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTT 276

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
Db        277 CGGAAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAGC 336

Qy        528 TTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG 587
Db        337 TTTAAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGG 396

Qy        588 CTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG 647
Db        397 GTTACTGAAGATTACCAAGCAGTGAATGTCTCCGTGAAAATTAATATATTGAGATCAGAG 456

Qy        648 GAAG 651
Db        457 AATG 460
```

RESULT 11
AV817525/c

```
LOCUS       AV817525                391 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  AV817525 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-94-M06 3',
            mRNA sequence.
ACCESSION   AV817525
```

```
VERSION     AV817525.1  GI:19859415
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 391)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
            and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .391
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL09-94-M06"
                     /clone_lib="RAFL9"
                     /dev_stage="plants at various developmental stages from
                     germination to mature seeds"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                     hr) treatments"
BASE COUNT      102 a    100 c     73 g    116 t
ORIGIN

  Query Match             11.5%;  Score 197.8;  DB 10;  Length 391;
  Best Local Similarity   71.7%;  Pred. No. 1.5e-33;
  Matches  259;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy       1051 GAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCG 1110
              |||  ||||| ||||| || || |||||| | || ||||||||||| || ||||||||
Db        388 GAGCTCAGCCCATGGCTCTAGAACAAGTACTTTCGGACCGGGATAGTGAAGATGAAGTTG 329

Qy       1111 ATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATA 1170
              ||||.||||| |||||||||||||||| |   |||||| |||| || || ||||||| ||
Db        328 ATGATGATGTGGCAGATTTTGAAGATAGAAGGATGCTCGATGATTTCGTTGATGTGACTA 269

Qy       1171 AAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAG 1230
              ||||||| || ||  | |||||  | |||||||||||||| || || ||  | ||  |||
Db        268 AAGATGAGAAACAGATGATGCACATGTGGAACTCGTTTGTGAGGAAGCAGCGAGTATTAG 209

Qy       1231 CAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGC 1290
              |||||||||| ||  | |||||||| || ||||| ||||||||  |||          ||
Db        208 CAGATGGTCACATTCCATGGGCATGCGAGGCATTCTCAAGATTGCACGGACCCATCATGG 149

Qy       1291 ACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATG 1350
                ||     |   | | ||||| |||||| ||||| || | ||||| ||||||||
Db        148 TTCGAACACCGCACTTGATTTGGTGCTGGAGAGTGTTTATGGTGAAACTGTGGAACCACG 89

Qy       1351 GACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATA 1410
              | ||| | ||  |    ||||| |||||||| |||||| | |||||  |   ||  | |
Db         88 GTCTTCTTGATGCCCGAACCATGAACAACTGTAATACCTTTCTCGAACAGCTCCAAATTT 29

Qy       1411 G 1411
              |
```

```
Db          28 G 28


RESULT 12
BQ740672
LOCUS       BQ740672                561 bp    mRNA    linear   EST 17-JUL-2002
DEFINITION  saq50e05.y1 Gm-c1076 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1076-3274 5' similar to TR:Q9ZNT9 Q9ZNT9
            FERTILIZATION-INDEPENDENT SEED 2 PROTEIN. ;, mRNA sequence.
ACCESSION   BQ740672
VERSION     BQ740672.1  GI:21887459
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 561)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact: ccu@resgen.com web site:
            www.resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .561
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1076-3274"
                     /clone_lib="Gm-c1076"
                     /tissue_type="wounded cotyledons"
                     /dev_stage="11 day old seedlings"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from mRNA isolated
                     from 11 day old seedlings treated with that were treated
                     with 2 ugs/ml of a crude glucan elicitor preparation
                     isolated from the mycelial walls of Phytophthora sojae.
                     The  library was prepared using the Stratagene pBluescript
                     II SK(+) library  construction kit. Complementary DNA was
                     synthesized from mRNA using a primer  consisting of a
                     poly(dT) sequence with an XhoI restriction site. EcoRI
                     adaptors were ligated to the blunt-ended cDNA fragments
                     followed by XhoI  digestion. The cDNA fragments were
                     directionally cloned into the EcoRI-XhoI  restriction site
                     of the pBluescript vector. The ligated cDNA fragments were
                     transformed into E.coli ElectroMax DH10B host cells. Plant
                     material was provided by Michael G. Hahn (Complex
                     Carbohydrate Research Center, University of Georgia) and
                     the library was constructed by Anu Khanna (Lila Vodkin lab
                     , University of Illinois)."
BASE COUNT      166 a     91 c    123 g    181 t
ORIGIN

  Query Match             11.0%;  Score 190;  DB 14;  Length 561;
  Best Local Similarity   72.2%;  Pred. No. 8.6e-32;
  Matches  247;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;
```

```
Qy       1068 CTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGAT 1127
Db          2 CTAGAACAAGTGTTATCAGACCGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGAT 61

Qy       1128 TTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTC 1187
Db         62 CTTGAAGATAGAAGGATGCTTGACGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTC 121

Qy       1188 ATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCT 1247
Db        122 ATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGGGTGCTAGCAGATGGTCATGTTCCG 181

Qy       1248 TGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTC 1307
Db        182 TGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAAGAGCTGATCTCATCTCCAGCTTTA 241

Qy       1308 TTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCC 1367
Db        242 TTTTGGTGTTGGAGGTTATTCATGATCAAACTTTGGAATCATGGTCTTCTTGATGCCTGT 301

Qy       1368 ACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAAT 1409
Db        302 ACAATGAACAACTGTAGCATAGTATTAGATAGTTACAGAAAT 343
```

```
RESULT 13
BG580495
LOCUS       BG580495                467 bp    mRNA    linear   EST 11-APR-2001
DEFINITION  EST482220 GVN Medicago truncatula cDNA clone pGVN-57B2 5' end, mRNA
            sequence.
ACCESSION   BG580495
VERSION     BG580495.1  GI:13595559
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town
            ,C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.
  TITLE     ESTs from one month old nitrogen-fixing root nodules of Medicago
            truncatula, 2001
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Carroll P. Vance
            Department of Agronomy and Plant Genetics
            University of Minnesota
            411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
            Tel: 612 625 5715
            Fax: 651-649-5058
            Email: vance004@maroon.tc.umn.edu
            University of Minnesota name: M381546e TIGR sequence name:
            MTCCT01TK More information is available at: http://www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Medicago truncatula"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pGVN-57B2"
                     /clone_lib="GVN"
                     /tissue_type="N2-fixing root nodules"
                     /dev_stage="effective root nodules harvested one month
                     post inoculation with Sinorhizobium meliloti"
                     /lab_host="E. coli strain XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     effective root nodules harvested one month post
                     inoculation with Sinorhizobium meliloti. The cDNA was
                     directionally ligated into the Uni-ZAP XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-ZAP phage using Ex-Assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      136 a     87 c     95 g    149 t
ORIGIN
```

```
  Query Match             11.0%;  Score 188.8;  DB 12;  Length 467;
  Best Local Similarity   67.0%;  Pred. No. 1.5e-31;
  Matches  300;  Conservative    0;  Mismatches  142;  Indels    6;  Gaps    2;

Qy        277 TTTCAACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTC 336
Db          1 TTACAGCTGATGAAAGTCTTTTGATTTATTGCAAACCTGTTGAATTCTACAATATTCTCT 60

Qy        337 ACCTTCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAG 396
Db         61 ATCGTCGTTCTCTTCACAATCCTACTTTTCTTAAGAGATGTTTGCGCTATAAAATAAAAG 120

Qy        397 CAAAGCGCAAAAGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTA 456
Db        121 CAAAGCAAAAAAGGAGGTTGCGA---GCAGGAATTGTGGTTTTCAACTATAGGGACTGTT 177

Qy        457 ATAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGC 516
Db        178 ACAATGCGCTTCGAAAGACTGAAGTGACTGAAGACTTTTCTTGTCCATTTTGCTTGATGC 237

Qy        517 TATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAAT 576
Db        238 AGTGTGCGAGCTTTAAGGGTTTGCGATTTCATCTTTGTGCATCACATGATCTATTCAACT 297

Qy        577 TTGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCT 636
Db        298 TTGAGTTCTGGGTCACCAATGATTACCAAGCAGTGAACGTCTCTGTAAAAACTGACATAT 357

Qy        637 TCATATTTGAGGAAGAAGGAAGTGATGACGAT---AAATTTGAGCCCTTCTCTCTCTGCT 693
Db        358 TGAGATCAGAGAATGTTGCTGATGGAGTAATTCCACAATCACAAACCTTCTTCTTCTGTT 417

Qy        694 CGAAACCTCGTAAGCGGAGACAAAGAGG 721
Db        418 CAAGATCTCGAAAGCGTAGACCAAGAGG 445
```

```
RESULT 14
AI164598
LOCUS       AI164598                364 bp    mRNA    linear   EST 03-DEC-1998
DEFINITION  A065P57U Hybrid aspen plasmid library Populus tremula x Populus
            tremuloides cDNA 5', mRNA sequence.
ACCESSION   AI164598
VERSION     AI164598.1  GI:3855883
KEYWORDS    EST.
SOURCE      Populus tremula x Populus tremuloides.
  ORGANISM  Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE   1  (bases 1 to 364)
  AUTHORS   Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A., Holmberg
            ,A., Amini,B., Bhalerao,R., Larsson,M., Villarroel,R., Van Montagu
            ,M., Sandberg,G., Olsson,O., Teeri,T.T., Boerjan,W., Gustafsson,P.,
            Uhlen,M., Sundberg,B. and Lundeberg,J.
  TITLE     Gene discovery in the wood-forming tissues of poplar: Analysis of
            5,692 expressed sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)
  MEDLINE   99007314
COMMENT     Contact: Sterky F
            Department of Biotechnology
            Royal Institute of Technology (KTH)
            Teknikringen 34, S-100 44 STOCKHOLM, Sweden
            Tel: +46 8 790 8287
            Fax: +46 8 24 54 52
            Email: fredrik@biochem.kth.se
            PCR PRimers
```

```
            FORWARD: AAAGGGGGATGTGCTGCAAGGCG
            BACKWARD: GCTTCCGGCTCGTATGTTGTGTG
            Seq primer: CGTTGTAAAACGACGGCCAG
            High quality sequence stop: 364.
FEATURES             Location/Qualifiers
     source          1. .364
                     /organism="Populus tremula x Populus tremuloides"
                     /db_xref="taxon:47664"
                     /clone_lib="Hybrid aspen plasmid library"
                     /tissue_type="Cambial region"
                     /dev_stage="1.5 m actively growing tree"
                     /lab_host="E.coli"
                     /note="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI;
                     Cambial region tissues, including developing xylem, the
                     meristematic  cambial zone and the developing and mature
                     phloem, was harvested  from 1.5 m actively growing trees.
                     cDNA was prepared and cloned  into lambda gt22a. DNA was
                     isolated and subcloned into pBluescript SK using SalI and
                     NotI restriction enzymes."
BASE COUNT      111 a     68 c     94 g     90 t      1 others
ORIGIN

  Query Match             10.2%;  Score 176.4;  DB 9;  Length 364;
  Best Local Similarity   75.5%;  Pred. No. 8.5e-29;
  Matches  244;  Conservative    0;  Mismatches   77;  Indels    2;  Gaps    2;

Qy        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010
              ||  || | |     | | ||||||||||| |||||   ||| ||  | ||    |||| |
Db         44 GCCATGCTACAATTTGCAAAGACAAGAAAATTATCAATTGAACGGNCTGACATGAGAAAC 103

Qy       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070
              |    ||| ||||| ||||| ||||| | ||| || || |||| ||||||||||   |
Db        104 CGTACACTCCTTCACAAACGACAATTTTTTCACTCACATAGAGCTCAGCCAATGGCAGCT 163

Qy       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130
              |||||||| ||||| || |||||||| |||||||||| || || ||||||||||||||||
Db        164 GAGCAAGTTATGTCAGATCGGGATAGTGAGGATGAAGTTGACGATGATGTTGCAGATTTT 223

Qy       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190
              ||||| ||   |||||||||||| ||||| |  |||| ||||||||| |||||| | |||
Db        224 GAAGACCGAAGGATGCTTGATGATTTTGTAG-CGTGACTAAAGATGAGAAGCAAATGATG 282

Qy       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250
              ||  | |||||||| |||||| || || ||  ||||  | |||||||| |||||| |  ||
Db        283 CACTTGTGGAACTCATTTGTGAGGAAGCAGCGGGTGCTTGCAGATGGACATATC-CCGGG 341

Qy       1251 GCATGTGAAGCATTTTCAAGATT 1273
              ||||| ||||||||  | |||||
Db        342 GCATGCGAAGCATTCACGAGATT 364


RESULT 15
BG314044
LOCUS       BG314044                405 bp    mRNA    linear   EST 16-APR-2001
DEFINITION  WHE2467_D04_G07ZS Triticum monococcum early reproductive apex cDNA
            library Triticum monococcum cDNA clone WHE2467_D04_G07, mRNA
            sequence.
ACCESSION   BG314044
VERSION     BG314044.1  GI:13115847
KEYWORDS    EST.
SOURCE      Triticum monococcum.
  ORGANISM  Triticum monococcum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
            ,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
            Stamova,B. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Early reproductive apex cDNA library from Triticum
            monococcum
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Triticum monococcum"
                     /cultivar="DV92"
                     /db_xref="taxon:4568"
                     /clone="WHE2467_D04_G07"
                     /clone_lib="Triticum monococcum early reproductive apex
                     cDNA library"
                     /tissue_type="Early reproductive apex"
                     /dev_stage="Seven week-old plants"
                     /lab_host="E. coli XLOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and
                     poly(A) RNA were prepared from apex at  double-ridge stage
                     to terminal-spikelet stage during transition from
                     vegetative state to flower state, a cDNA library was made,
                     and the  cDNA clones were in vivo excised at the
                     University of California, Davis  (V. Echenique, B. Stamova
                     , J. Dubcovsky).  Plasmid DNA preparations and  DNA
                     sequencing were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT      105 a     85 c    104 g    111 t
ORIGIN

  Query Match             10.1%;  Score 174.6;  DB 12;  Length 405;
  Best Local Similarity   68.8%;  Pred. No. 2.2e-28;
  Matches  240;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy       1056 CAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGAC 1115
              ||||||||||||  | || |||||  |||| ||     || ||||||||||||||| |||||
Db          4 CAGCCAATGGCAGTCGAACAAGTTTTGTCAGATGATGACAGCGAGGATGAAGTTGATGAT 63

Qy       1116 GATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGAT 1175
              ||  |||| || || |||||| | | |||||||| ||| ||| | |||||  |  ||||||
Db         64 GACATTGCCGACTTGGAAGATAGACGGATGCTTGCTGATTTTCTTGATGTCACGAAAGAT 123

Qy       1176 GAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGAT 1235
              || ||||    | |||||| | ||||| || ||| |   | ||||||||||  |||| |||
Db        124 GAGAAGCTTATTATGCATATGTGGAATTCATTTATTCGGAAACAAAGGGTGCTAGCTGAT 183

Qy       1236 GGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGT 1295
              || |||||  ||||||| ||||| |  || ||  |  || | |       || |    |
Db        184 GGGCATATACCTTGGGCCTGTGAGGGCTTCTCCCGGCTTCATGGACCGCAGCTTGTACAA 243

Qy       1296 TACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTT 1355
              || |  | ||  ||||||| ||| | ||| |  ||||||| || ||||||||  | ||
Db        244 AACCCTCCTCTGCTCTGGTGCTGGCGTTTGGTGATGATTAAGCTGTGGAACCACAGCCTG 303

Qy       1356 GTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCC 1404
              | ||  |   | |||| ||||  ||||| ||||| || |||   | ||
Db        304 CTGGATGCGCGCGCCATGAACACATGCAACACCATTCTTGAGGGATACC 352


Search completed: June 19, 2003, 22:08:41
Job time : 2572 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:10:48 ; Search time 1286 Seconds
(without alignments)
9309.105 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagcttcttcaattttgct.....aatggattctgcttataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7816032 seqs, 3476047932 residues

Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	366.8	21.3	1722	10	US-10-425-114-11348	Sequence 11348, A
	2	366.8	21.3	1838	10	US-10-424-599-111422	Sequence 111422,
	3	257.4	14.9	1841	10	US-10-425-114-31883	Sequence 31883, A
	4	236.2	13.7	2640	9	US-10-231-778-6	Sequence 6, Appli
c	5	230.2	13.4	1004	10	US-10-437-963-11609	Sequence 11609, A
c	6	227.2	13.2	1070	10	US-10-424-599-141143	Sequence 141143,
	7	192.4	11.2	716	10	US-10-424-599-111421	Sequence 111421,
	8	185.6	10.8	443	5	US-09-837-604B-13507	Sequence 13507, A
	9	185.6	10.8	443	7	US-09-837-604A-13507	Sequence 13507, A
	10	174	10.1	558	10	US-10-389-048-21629	Sequence 21629, A
	11	158.8	9.2	2700	10	US-10-437-963-91042	Sequence 91042, A
	12	138.8	8.1	554	5	US-09-531-113-5697	Sequence 5697, Ap
	13	138.8	8.1	554	6	US-09-531-113-5697	Sequence 5697, Ap
c	14	125.6	7.3	322	10	US-10-424-599-40645	Sequence 40645, A
	15	104	6.0	453	5	US-09-837-604B-6414	Sequence 6414, Ap
	16	104	6.0	453	7	US-09-837-604A-6414	Sequence 6414, Ap
	17	104	6.0	453	10	US-10-437-963-97425	Sequence 97425, A
	18	101.2	5.9	2439	10	US-10-425-114-30709	Sequence 30709, A
c	19	100	5.8	531	5	US-09-837-604B-940	Sequence 940, App
c	20	100	5.8	531	7	US-09-837-604A-940	Sequence 940, App
c	21	91.8	5.3	632	10	US-10-424-599-73456	Sequence 73456, A
	22	82.8	4.8	250	5	US-09-531-113-40606	Sequence 40606, A
	23	82.8	4.8	250	6	US-09-531-113-40606	Sequence 40606, A
	24	71	4.1	274	10	US-10-389-048-16407	Sequence 16407, A
	25	69	4.0	921	10	US-10-424-599-108137	Sequence 108137,
c	26	63.8	3.7	201	13	US-60-466-412-340472	Sequence 340472,
c	27	63.8	3.7	56447	13	US-60-466-412-86674	Sequence 86674, A
	28	58.2	3.4	939	10	US-10-437-963-81637	Sequence 81637, A
	29	57.8	3.4	6458	9	US-10-231-778-7	Sequence 7, Appli
	30	57.2	3.3	241	10	US-10-424-599-52847	Sequence 52847, A
c	31	57.2	3.3	492	9	US-10-144-771-38577	Sequence 38577, A
	32	56.8	3.3	961	9	US-10-144-771-31796	Sequence 31796, A
c	33	54.2	3.1	8524	9	US-10-311-455-1814	Sequence 1814, Ap
	34	53.2	3.1	2892	1	PCT-US01-29288A-255	Sequence 255, App
	35	53.2	3.1	2892	10	US-10-149-310-255	Sequence 255, App
	36	53	3.1	350	10	US-10-424-599-9668	Sequence 9668, Ap
c	37	52.2	3.0	8662	9	US-10-240-485-190	Sequence 190, App
	38	51.6	3.0	867	9	US-10-321-856-340	Sequence 340, App
	39	51.6	3.0	867	9	US-10-321-856-342	Sequence 342, App
	40	51.6	3.0	1397	9	US-10-321-856-343	Sequence 343, App
c	41	51.6	3.0	1397	9	US-10-321-856-345	Sequence 345, App
	42	51.4	3.0	958	9	US-10-144-771-24706	Sequence 24706, A
	43	51.4	3.0	1178	10	US-10-425-114-16009	Sequence 16009, A
c	44	50.8	3.0	1141	6	US-09-806-708B-22	Sequence 22, Appl
c	45	50.6	2.9	3673778	9	US-10-312-841-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-425-114-11348
; Sequence 11348, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11348
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701055296_FLI
US-10-425-114-11348

Query Match 21.3%; Score 366.8; DB 10; Length 1722;
Best Local Similarity 59.8%; Pred. No. 1.1e-80;
Matches 757; Conservative 0; Mismatches 422; Indels 87; Gaps 5;

```
Qy      228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
            || |||||  |||| |||| ||  |   |     |    || ||||  ||| || |||||
Db      161 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 220

Qy      288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
            ||||  || ||||| ||||| || ||||||  ||| || || ||  |  |||  ||  ||
Db      221 GAGAGCCTTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCT 280

Qy      348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
            ||    || || || ||||||   ||||| |||   || |||||  ||||||  || |||
Db      281 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAA 340

Qy      408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
            || | ||  |||    | || || ||  ||||||| |||| |||  |  | ||||   |
Db      341 AGGAGGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTT 397
```

```
Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
Db        398 CGAAAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAAC 457

Qy        528 TTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG 587
Db        458 TTTAAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGG 517

Qy        588 CTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG 647
Db        518 GTTACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGATATATTGAGATCAGAG 577

Qy        648 GAAGAAGGAAGTGATGACGAT---AAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGT 704
Db        578 AATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCTCGA 637

Qy        705 AAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCG 764
Db        638 AAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTGGAG 697

Qy        765 TTGGATTCACCCA---------------------------------------------- 777
Db        698 TTGGACTCACCAGAAGGCATACATAATGGATTTCTACAAAAAGATGATGATATCTTATCC 757

Qy        778 -----------------GTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGA 820
Db        758 TGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTCCCTAGTGGAAGA 817

Qy        821 T--------GGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTT 872
Db        818 AATGACGGAGGAAAATTTGGTCCTGATCATCCTGGCACCATGGACAACCTGGAGCATGTG 877

Qy        873 GAGATGACCAGCAACATTCCACCA---------GCCATAGCCCACTCTTCTCTGGACGCT 923
Db        878 GAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCTTCTGTGGACCCT 937

Qy        924 GGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTA 983
Db        938 GAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCTAAAACAAAGAAGCTA 997

Qy        984 TCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCAT 1043
Db        998 AGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAGAGACTTTTCTTTCAC 1057

Qy       1044 TCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGAT 1103
Db       1058 TCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGACCGTGATAGTGAAGAC 1117

Qy       1104 GAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGAT 1163
Db       1118 GAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTTGACGATTTTGTGGAT 1177

Qy       1164 GTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGG 1223
Db       1178 GTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGG 1237

Qy       1224 GTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAA 1283
Db       1238 GTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAA 1297

Qy       1284 GAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1343
Db       1298 GAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTCATGATCAAACTTTGG 1357

Qy       1344 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGC 1403
Db       1358 AATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATAGTATTAGATAGTTAC 1417

Qy       1404 CGTAAT 1409
Db       1418 AGAAAT 1423
```

```
RESULT 2
US-10-424-599-111422
; Sequence 111422, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111422
;   LENGTH: 1838
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1
US-10-424-599-111422

  Query Match             21.3%;  Score 366.8;  DB 10;  Length 1838;
  Best Local Similarity   59.8%;  Pred. No. 1.1e-80;
  Matches  757;  Conservative    0;  Mismatches  422;  Indels   87;  Gaps    5;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
Db        255 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 314

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
Db        315 GAGAGCCTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCT 374

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
Db        375 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAA 434

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
Db        435 AGGAGGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTT 491

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
Db        492 CGAAAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAAC 551

Qy        528 TTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG 587
Db        552 TTTAAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGG 611

Qy        588 CTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG 647
Db        612 GTTACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGATATATTGAGATCAGAG 671

Qy        648 GAAGAAGGAAGTGATGACGAT---AAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGT 704
Db        672 AATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCTCGA 731

Qy        705 AAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCG 764
Db        732 AAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTGGAG 791

Qy        765 TTGGATTCACCCA---------------------------------------------- 777
Db        792 TTGGACTCACCAGAAGGCATACATAATGGATTTCTACAAAAAGATGATGATATCTTATCC 851

Qy        778 -----------------GTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGA 820
Db        852 TGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTCCCTAGTGGAAGA 911

Qy        821 T--------GGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTT 872
```

```
Db        912 AATGACGGAGGAAAATTTGGTCCTGATCATCCTGGCACCATGGACAACCTGGAGCATGTG 971

Qy        873 GAGATGACCAGCAACATTCCACCA---------GCCATAGCCCACTCTTCTCTGGACGCT 923
              ||    |    ||||||||||           |||||  |||| ||||||  |||||  ||
Db        972 GAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCTTCTGTGGACCCT 1031

Qy        924 GGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTA 983
              |    ||      | |    |     | |   |  |   || |||| ||||  |||  ||
Db       1032 GAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCTAAAACAAAGAAGCTA 1091

Qy        984 TCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCAT 1043
                 || ||||||||  | |||| ||   | |||||  |||||  | |  |||| |||
Db       1092 AGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAGAGACTTTTCTTTCAC 1151

Qy       1044 TCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGAT 1103
              || ||||||||||||||  |||||  || ||  |||||  | ||  |||||  |||||  || ||
Db       1152 TCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGACCGTGATAGTGAAGAC 1211

Qy       1104 GAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGAT 1163
              |||||  ||||||||  ||||||||  ||||||||  |   |||||||||  ||  |||||||||
Db       1212 GAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTTGACGATTTTGTGGAT 1271

Qy       1164 GTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGG 1223
              ||    ||||||||||  ||  ||||||||||  ||||||||  |||  | ||  ||  ||||||
Db       1272 GTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTTATGAGGAAGCAAAGG 1331

Qy       1224 GTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAA 1283
              ||  ||||||||||||||  |  | |||||  |||||  ||||||||  |    ||  |  |||
Db       1332 GTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCCAAGCTTCATGGAAAA 1391

Qy       1284 GAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1343
              |||  ||  |    |   ||  |  | ||  ||||||||||  ||  ||   ||||  |||||  |||
Db       1392 GAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTCATGATCAAACTTTGG 1451

Qy       1344 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGC 1403
              ||  |||||  |||  |  ||   |     ||  ||  ||||||||  |   |   |  ||  |  ||  |
Db       1452 AATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATAGTATTAGATAGTTAC 1511

Qy       1404 CGTAAT 1409
                 | |||
Db       1512 AGAAAT 1517


RESULT 3
US-10-425-114-31883
; Sequence 31883, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31883
;   LENGTH: 1841
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12_FLI
US-10-425-114-31883

  Query Match            14.9%;  Score 257.4;  DB 10;  Length 1841;
  Best Local Similarity   56.5%;  Pred. No. 1.8e-53;
  Matches  573;  Conservative    0;  Mismatches  391;  Indels   51;  Gaps    3;

Qy        424 CTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTA 483
              ||||  |  ||| |||| || || ||| | |  || || ||| | || || ||||||||||

Db        493 CTGGAAATGTACTTTTTAATTACAAGTACTACAACAATACAATGCAAAAGACTGAAGTTA 552

Qy        484 GGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAAT 543
                || |||| ||| || ||||||||||   | | |||||| ||||||||||| ||   ||
Db        553 CTGAAGATTTTTCCTGCCCATTTTGCTTGGTACCATGTGGAAGCTTCAAGGGCCTAGGAT 612

Qy        544 TTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACC 603
              || || |||||||| |||||  | ||  | | ||||||   | | |  |||||  | ||
Db        613 GCCACTTAAATTCATCACATGACCTGTTCCACTATGAGTTGTGGATATCTGAAGAGTGCC 672

Qy        604 AGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAAGAAGGAAGTGATG 663
              || | || ||||||  | | ||   | || |||  | |  ||||   |  | |   |  |
Db        673 AGGCTGTGAATGTTAGTCTGAAGGCTGATGCCTGGAAAACTGAGTTTGTGGCAGAGGGAG 732

Qy        664 ACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTG 723
                |||  |   | |     | |  || ||          ||||  | |  |||   |
Db        733 TTGATCCAAGGCATCAAACATTTTCCTACTGCTCGAGGTTTAAGAAGCGTAGAAG----G 788

Qy        724 GCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCCAGTTTAA 783
                ||||  ||| | |   | || | | ||| |       | |   ||  |  |||  |
Db        789 TTGGAAACCACAGCTGAGAAATTCAGGCATGTTCACCCACATATTATGGAATCAGGAT-- 846

Qy        784 CTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTAGGATATC 843
                                              || | | ||||   |    || |  | | ||
Db        847 --------------------------CACATGAAGATGCCCAGGCAGGGTCTGAAGATG 879

Qy        844 CCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCAGCCATAG 903
                | | |   |  |    |||  |  | ||                  | ||   |
Db        880 ACTATGTGCAAAGGGAAAATGGGCTTTCTG------------------TAGCAAATACTT 921

Qy        904 CCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTG 963
              |         ||   |  |||    |      |   ||     | | |||  | |
Db        922 CAGTTGACCCTAGTCAGCCTGTACATGGGAGCAATCTTTCACCACCAACAGTGCTACAGT 981

Qy        964 CTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTC 1023
              |   |||  |||||||  |||||||  ||||||||  ||   | ||||  ||   | |||   | |
Db        982 TTGGAAAGTCAAGAAAACTATCTGCCGAGCGATCTGACCCCAGAAACCGACAACTGTTGC 1041

Qy       1024 AGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGT 1083
              | ||  ||  ||  ||||  |||||  |||||  |  |||||||||  |   |  ||  |||||   | |
Db       1042 AAAAGCGTCAGTTCTTCCATTCACACAGGGCACAGCCAATGCCATTGGAACAAGTCCTCT 1101

Qy       1084 CTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGA 1143
              | ||  |  ||||||||  ||  ||  ||||||  ||   ||||  ||  |||||  |||  |  ||
Db       1102 CAGATCATGATAGCGAAGACGAGGTTGATGATGACATTGCCGACTTTGAGGATAGACGGA 1161

Qy       1144 TGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACT 1203
              ||||  ||  ||  |||||  |||||  |  ||||||||  ||||    |  ||||||  |  |||||  |
Db       1162 TGCTCGACGATTTTGTTGATGTTACAAAAGATGAGAAGCTTATTATGCATATGTGGAATT 1221

Qy       1204 CGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAAGCAT 1263
              | |||||  ||||  ||||||||  ||||  ||||||||||  |  ||||||||  ||  ||  |
Db       1222 CATTTGTTCGAAAGCAAAGGGTGCTAGCCGATGGTCATATACCGTGGGCATGCGAGGCGT 1281

Qy       1264 TTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGAT 1323
              | ||  ||||    |  |   ||  ||  |     |    ||  |  ||  |  ||| |  ||| | |
Db       1282 TCTCCAGATGCCATGGACAACAGCTTGTACAAAACCCTGCCCTGCTGTGGGGCTGGCGTT 1341

Qy       1324 TGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCA 1383
              | ||  ||||||||||  ||||||||  | |||  | |||  |   | ||||  ||    ||||
Db       1342 TCTTCATGATTAAACTCTGGAACCACAGCCTTCTTGACGCCCGCGCCATGAATGCTTGCA 1401

Qy       1384 ATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAACAACAA 1438
              | ||  |||||  |    | ||   |  |    |    | |||     | | ||||
Db       1402 ACACAATCCTTCAAGGCTACCAAGACGGAAGTGCGGACCCCAAGTAAATCTACAA 1456


RESULT 4
US-10-231-778-6
; Sequence 6, Application US/10231778
; GENERAL INFORMATION:
;  APPLICANT: Bilodeau, Pierre
;  APPLICANT:  Chaudhury, Abdul M.
```

```
;  APPLICANT:  Dennis, Elizabeth S.
;  APPLICANT:  Koltunow, Anna M.G.
;  APPLICANT:  Luo, Ming
;  APPLICANT:  Peacock, William J.
;  TITLE OF INVENTION: Method for inducing seed development by down-regulating
;  TITLE OF INVENTION:  expression of the FIS2 gene
;  FILE REFERENCE: 72-98A
;  CURRENT APPLICATION NUMBER: US/10/231,778
;  CURRENT FILING DATE:  2002-11-08
;  PRIOR APPLICATION NUMBER: 09/398,237
;  PRIOR FILING DATE: 1999-09-20
;  PRIOR APPLICATION NUMBER: 60/101,184
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: AU PP6061
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6062
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6063
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PQ1345
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: AU PQ1346
;  PRIOR FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 239
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 2640
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2439)
;   OTHER INFORMATION: Nucleotides from 1 to 2439 represent protein
;   OTHER INFORMATION: coding sequence.
US-10-231-778-6

  Query Match             13.7%;  Score 236.2;  DB 9;  Length 2640;
  Best Local Similarity   71.0%;  Pred. No. 3.7e-48;
  Matches  313;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        954 GTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGCCAC 1013
              |||   | || || ||  | || | |||||  ||| |||||  |  ||||||| |||
Db       1915 GTGACTCGTGTTAGTAGAAGAAAAGAGTTACATGCAGAGCGGTGCGAGGCTAAAAGATTG 1974

Qy       1014 CTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTTGAG 1073
                | |||| ||   || || |||||||| || || | | | |||||||||  |  ||||
Db       1975 GAGCGTCTTAAGGGTCGACAGTTCTATCACTCCCAAACAATGCAGCCAATGACTTTTGAA 2034

Qy       1074 CAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTTGAA 1133
              |||||||||||| ||  |||||||||| ||||    |||||  ||| |  |||| ||
Db       2035 CAAGTAATGTCTAACGAGGATAGCGAGAATGAGACTGATGATTATGCTTTAGATATTAGC 2094

Qy       1134 GATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCAT 1193
              || |||| ||  |||||     |||||| |||||  |||||  ||||||| || |||| ||
Db       2095 GAACGCCTGAGACTTGAACGTCTTGTGGGTGTGAGCAAAGAGGAAAAGCGATACATGTAT 2154

Qy       1194 CTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGGGCA 1253
              |||||||||    |||||| |||||||||||||| || || ||||| ||| |  ||||||||
Db       2155 CTTTGGAACATATTTGTACGAAAACAAAGGGTGATCGCGGATGGACATGTTCCTTGGGCA 2214

Qy       1254 TGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGG 1313
              |||||||  ||| ||| | || |  || ||||| || |    ||  |||||  ||   |||
Db       2215 TGTGAAGAGTTTGCAAAACTTCATAAGGAAGAGATGAAGAATTCTTCATCTTTCGATTGG 2274

Qy       1314 TGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACCATC 1373
              || |||||| |||||  |||||||||| |||||| | || ||  ||  | |     ||| ||
Db       2275 TGGTGGAGAATGTTTAGGATTAAACTGTGGAACAACGGTCTCATCTGCGCCAAGACCTTC 2334

Qy       1374 AACAACTGCAATACCATCCTC 1394
                |||| |||| ||||||||||
Db       2335 CACAAATGCACTACCATCCTC 2355
```

```
RESULT 5
US-10-437-963-11609/c
; Sequence 11609, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 11609
;   LENGTH: 1004
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609

  Query Match             13.4%;  Score 230.2;  DB 10;  Length 1004;
  Best Local Similarity   68.8%;  Pred. No. 8.6e-47;
  Matches  316;  Conservative    0;  Mismatches  143;  Indels    0;  Gaps    0;

Qy        950 AGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAG 1009
              | | ||| | |    |  ||||||||||||| | |||| ||| ||| | ||  | ||||
Db        782 AACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGAAA 723

Qy       1010 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCT 1069
                |  |  || || || |||||||| |||| |||||||||||||  |  || ||||||||  |
Db        722 TCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCATT 663

Qy       1070 TGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTT 1129
              ||||||||  | || || || ||||| || |||||||| ||||| ||  |||| ||||||
Db        662 GGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGATTT 603

Qy       1130 TGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCAT 1189
              ||||||| |  |||||||||| ||||| ||||| |  ||||| || || |  | ||
Db        602 TGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATTAT 543

Qy       1190 GCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTG 1249
              |||| | ||||| || |||||  | ||||||||||  |||| ||||| |||||  | ||
Db        542 GCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCCTG 483

Qy       1250 GGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTT 1309
              |||||| |||||||| ||   ||| | |  ||||  |  |  |  || | ||  |
Db        482 GGCATGCGAAGCATTCTCGCAGTTTCATGGACAAGAACTTGTACAAAATCCAGCTCTACT 423

Qy       1310 CTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCAC 1369
              |||||||||||| || ||| || | ||||| ||||||||  | ||  | ||  |     |
Db        422 ATGGTGTTGGAGGTTTTTTATGGTCAAACTCTGGAACCACAGTCTACTGGATGCGCGAGC 363

Qy       1370 CATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAA 1408
              ||| ||   |||||| || || || ||    | ||  ||
Db        362 CATGAATGCCTGCAACACAATTCTTGAAGGCTACCTGAA 324


RESULT 6
US-10-424-599-141143/c
; Sequence 141143, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
```

```
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141143
;   LENGTH: 1070
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1
US-10-424-599-141143

  Query Match            13.2%;  Score 227.2;  DB 10;  Length 1070;
  Best Local Similarity  68.6%;  Pred. No. 4.9e-46;
  Matches  313;  Conservative    0;  Mismatches  143;  Indels    0;  Gaps    0;

Qy        949 AAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAA 1008
              |||| ||| | |    | | ||||||| ||||| ||   ||||||  | ||  |  |||
Db        824 AAGCAGTGCTTCAGTTTGCCAAGACAAGGAAGTTGTCAATTGAGCGTCCTGACCCACGAA 765

Qy       1009 GCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGC 1068
                |      |  |   ||| || ||||| | |||||| |||| ||  ||||||||||
Db        764 ACAGTACCTTCTTGAGGAAGCGACAATTTTTTCATTCACACAAAGCTCAGCCAATGGCAA 705

Qy       1069 TTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATT 1128
              |||| |||||  | || ||    |||||||| |||||||| ||||| ||||||||| ||||
Db        704 TTGAACAAGTTCTATCCGATAAAGATAGCGAAGATGAAGTTGATGATGATGTTGCCGATT 645

Qy       1129 TTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCA 1188
              |||||||||   |||||||||  |  |||| |||||||  || ||||| |||   ||||
Db        644 TTGAAGATCGAAGGATGCTTGAAAATGTTGTTGATGTGAGCAATGATGAGAAGACTTTCA 585

Qy       1189 TGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTT 1248
              ||||| | |||||||| |||||  | || ||  | || || ||||||||||| || || |
Db        584 TGCATATGTGGAACTCATTTGTTCGGAAGCATCGTGTGATTGCAGATGGTCACATTTCAT 525

Qy       1249 GGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCT 1308
              |||||||||| || || |||| |||  | |     |||||    |  |   | |||||
Db        524 GGGCATGTGAGGCTTTCTCAAAATTGCATGCACCTGAGTTTGTTCAATCTCCCTCACTGG 465

Qy       1309 TCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCA 1368
                ||||||||||| | ||| || | ||| |||  || |||||  ||| | ||  |    |
Db        464 CAGGGTGTTGGAGAATATTTATGGTCAAATTATACAATCATGGTCTTCTAGATGCTCGGA 405

Qy       1369 CCATCAACAACTGCAATACCATCCTCGAGAATTGCC 1404
              ||||  ||  |||| ||||  || || ||| | | ||
Db        404 CCATGAATGACTGTAATATTATTCTTGAGCAATACC 369


RESULT 7
US-10-424-599-111421
; Sequence 111421, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111421
;   LENGTH: 716
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1
US-10-424-599-111421

  Query Match            11.2%;  Score 192.4;  DB 10;  Length 716;
  Best Local Similarity  68.5%;  Pred. No. 2e-37;
  Matches  281;  Conservative    0;  Mismatches  126;  Indels    3;  Gaps    1;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
              || |||||  |||| |||| ||  |   |     |    || ||||  ||| || |||||
Db        214 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 273

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
              |||| ||| || || ||||| || ||||||  ||| || || ||  |  |||  ||| ||
Db        274 GAGAGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCT 333

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
              ||    || || || ||||||   ||||| |||   || |||||  |||||| ||| |||
Db        334 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAA 393

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
              || | ||  |||    | || || ||  |||||| |||| ||||    | ||||   |
Db        394 AGGAGGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTT 450

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
              | |||||||||||| |  || || | |||||||| |||||||  ||||  ||||| |||
Db        451 CGGAAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAGC 510

Qy        528 TTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG 587
              || |||||  ||| ||||||| |   |||||| |||||| |||  | |||| ||||  |
Db        511 TTTAAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGGGTTCTGG 570

Qy        588 CTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTT 637
                ||   ||||| ||||| |||| ||||| || || ||| |||||   ||
Db        571 GTTACTGAAGATTACCAAGCAGTGAATGTCTCCGTGAAAATTAATATATT 620


RESULT 8
US-09-837-604B-13507
; Sequence 13507, Application US/09837604B
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  APPLICANT:  De La Pena, Robert C.
;  APPLICANT:  Bougri, Olegs
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(51892)B
;  CURRENT APPLICATION NUMBER: US/09/837,604B
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: US 60/197,872
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
;   LENGTH: 443
;   TYPE: DNA
;   ORGANISM: Oryza sativa nipponbare
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604B-13507

  Query Match            10.8%;  Score 185.6;  DB 5;  Length 443;
  Best Local Similarity  72.0%;  Pred. No. 8.3e-36;
  Matches  242;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        950 AGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAG 1009
              | | ||| | |    |  ||||||||||| | |||| ||| ||| | ||  | ||||
Db         98 AACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGAAA 157

Qy       1010 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCT 1069
                |  |  || || || |||||||| |||| |||||||||||| |  || ||||||||  |
Db        158 TCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCATT 217

Qy       1070 TGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTT 1129
                ||||||||  | || || || ||||| || |||||||| ||||| ||  |||| |||||
Db        218 GGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGATTT 277

Qy       1130 TGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCAT 1189
```

```
Db        278 TGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATTAT 337

Qy       1190 GCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTG 1249
Db        338 GCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCCTG 397

Qy       1250 GGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGA 1285
Db        398 GGCATGCGAAGCATTTCTCGAGTTTCATGGACAAGA 433


RESULT 9
US-09-837-604A-13507
; Sequence 13507, Application US/09837604A
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  APPLICANT:  De La Pena, Robert C.
;  APPLICANT:  Bougri, Olegs
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(51892)B
;  CURRENT APPLICATION NUMBER: US/09/837,604A
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: US 60/197,872
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
;   LENGTH: 443
;   TYPE: DNA
;   ORGANISM: Oryza sativa nipponbare
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604A-13507

  Query Match             10.8%;  Score 185.6;  DB 7;  Length 443;
  Best Local Similarity   72.0%;  Pred. No. 8.3e-36;
  Matches  242;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        950 AGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAG 1009
Db         98 AACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGAAA 157

Qy       1010 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCT 1069
Db        158 TCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCATT 217

Qy       1070 TGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTT 1129
Db        218 GGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGATTT 277

Qy       1130 TGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCAT 1189
Db        278 TGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATTAT 337

Qy       1190 GCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTG 1249
Db        338 GCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCCTG 397

Qy       1250 GGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGA 1285
Db        398 GGCATGCGAAGCATTTCTCGAGTTTCATGGACAAGA 433


RESULT 10
US-10-389-048-21629
; Sequence 21629, Application US/10389048
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka
;  APPLICANT:  Shenk, Michael A.
;  TITLE OF INVENTION: Polynucleotides, Materials Incorporating
;  TITLE OF INVENTION:  Them, and Methods for Using Them
;  FILE REFERENCE: 11000.1041U1c1
;  CURRENT APPLICATION NUMBER: US/10/389,048
;  CURRENT FILING DATE:  2003-03-13
;  NUMBER OF SEQ ID NOS: 25129
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21629
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-10-389-048-21629

  Query Match             10.1%;  Score 174;  DB 10;  Length 558;
  Best Local Similarity   62.8%;  Pred. No. 6.8e-33;
  Matches  304;  Conservative    0;  Mismatches  175;  Indels    5;  Gaps    2;

Qy        221 TTGAACAAGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTC 280
Db         64 TTGGGTTACAATGTGCCATCAAAGTTCTTGTGAACATTTTTCTGATGAAGATGCAACTGC 123

Qy        281 AACTGATGAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCT 340
Db        124 AGCAGAAGAGAATCTCCTGGTGTATTGCAAGCCTGTTGAGCTATACAACATTCTACACCG 183

Qy        341 TCGCTCTCTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAA 400
Db        184 CCGTGCCATACATAATCCTTCTTTCCTCCGGAGATGTTTGCGTTTCAAAATACAAGCAAG 243

Qy        401 GCGCAAAAGAAAGTCAAGATCTA---CTGGGATGGTAGTTTTCAACTATAAGGATTGTAA 457
Db        244 ACGTAAGCGTAATAAGAGGTTGAGGGCTGGAATAGTCATTTTCAACTATAGGGACCGCAA 303

Qy        458 TAACACATTACAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCT 517
Db        304 CAACGTGCATCAAAAGGTTGAAGTAACTGAAGATTTCTCTTGCCCATTTTGCTTAATGCA 363

Qy        518 ATGTGGTAGCTTCAAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATT 577
Db        364 ATGTGCAAGCTTTAAGGGTCTAAGGTTGCATTTATGCTCATCGCATGACTTGTTCAACTT 423

Qy        578 TGAGTTCAAGCTTTTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTT 637
Db        424 CGAGTTTTGGGTTATTGAAGAGTATCAGGCAGTGAATGTCTCTGTTAAAGTTGATATGCT 483

Qy        638 CATATTTGAGG--AAGAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCG 695
Db        484 GAGATCCGAGGTTCTGACTGAAGAAGCAACCCTAAATTGTGGACATTCTTCTTCTGCCCA 543

Qy        696 AAAC 699
Db        544 AACC 547


RESULT 11
US-10-437-963-91042
; Sequence 91042, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91042
;   LENGTH: 2700
```

```
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_89656C.1
US-10-437-963-91042

  Query Match             9.2%;  Score 158.8;  DB 10;  Length 2700;
  Best Local Similarity   66.9%;  Pred. No. 6.6e-29;
  Matches  226;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy       1059 CCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGAT 1118
             |||||| |  | ||  |||| | || ||    ||||| || |||||||| ||||| |||
Db       1983 CCAATGACATTCGAAGAAGTTCTCTCAGATAATGATAGTGAAGATGAAGTAGATGATGAT 2042

Qy       1119 GTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAA 1178
              |||| ||||| |||||| |   ||||||||||||| ||||| ||||| |  ||||||||
Db       2043 ATTGCTGATTTGGAAGATAGAAGGATGCTTGATGATTTTGTTGATGTTACAAAAGATGAG 2102

Qy       1179 AAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGT 1238
              ||||  | |||||  | ||||| || ||| | ||||||||||  |  |||| ||| ||
Db       2103 AAGCGCATTATGCACATGTGGAATTCATTTATTCGAAAACAAAGTATACTAGCTGATAGT 2162

Qy       1239 CATATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTAC 1298
              ||  |  ||||||| ||||| ||||| ||  ||  | | |   |||| |        ||
Db       2163 CACGTACCTTGGGCTTGTGAGGCATTCTCCCGACATCATGGAGAAGAACTTTTAGAAAAC 2222

Qy       1299 TCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTC 1358
              ||  |  |  | ||| | ||| |  ||||| |||| ||||| ||||| ||  | ||  ||
Db       2223 TCCGCTTTACTATGGGGATGGCGCATGTTTATGATCAAACTCTGGAATCACAGTCTACTC 2282

Qy       1359 GACTCAGCCACCATCAACAACTGCAATACCATCCTCGA 1396
                |  ||| ||  ||| |||||| |  || || ||
Db       2283 TCTGCCCGCACAATGGACACCTGCAACAGAATTCTTGA 2320


RESULT 12
US-09-531-113-5697
; Sequence 5697, Application US/09531113
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  La Rosa, Thomas J.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15761)B
;  CURRENT APPLICATION NUMBER: US/09/531,113
;  CURRENT FILING DATE:  2000-03-22
;  NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

  Query Match             8.1%;  Score 138.8;  DB 5;  Length 554;
  Best Local Similarity   67.0%;  Pred. No. 3.9e-24;
  Matches  213;  Conservative    0;  Mismatches  102;  Indels    3;  Gaps    1;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
              || |||||  |||| |||| ||  |   |     |  || ||||  ||| || ||||||
Db        190 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 249

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
              |||| ||| || || ||||| || ||||||  ||| || || ||  |  |||  ||| ||
Db        250 GAGAGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCT 309

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
              ||    || || || ||||||   ||||| |||   || |||||  |||||| ||| |||
Db        310 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAA 369

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
              || ||||  |||    | || || ||  ||||||| |||| ||||    | ||||   |
Db        370 AGGAAGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTT 426

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
              | ||||||||||| |  || || | |||||||| |||||||  ||||  ||||| |||
Db        427 CGGAAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGCAGC 486

Qy        528 TTCAAGGGGCTGCAATTT 545
              || | |   ||| |  ||
Db        487 TTTAGGTTACTGGAGATT 504


RESULT 13
US-09-531-113-5697
; Sequence 5697, Application US/09531113
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  La Rosa, Thomas J.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15761)B
;  CURRENT APPLICATION NUMBER: US/09/531,113
;  CURRENT FILING DATE:  2000-03-22
;  NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

  Query Match             8.1%;  Score 138.8;  DB 6;  Length 554;
  Best Local Similarity   67.0%;  Pred. No. 3.9e-24;
  Matches  213;  Conservative    0;  Mismatches  102;  Indels    3;  Gaps    1;

Qy        228 AGAATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGAT 287
              || |||||  |||| |||| ||  |   |     |  || ||||  ||| || ||||||
Db        190 AGCATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGAT 249

Qy        288 GAGAATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCT 347
              |||| ||| || || ||||| || ||||||  ||| || || ||  |  |||  ||| ||
Db        250 GAGAGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCT 309

Qy        348 CTAGGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAA 407
              ||    || || || ||||||   ||||| |||   || |||||  |||||| ||| |||
Db        310 CTTCAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAA 369

Qy        408 AGAAAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTA 467
              || ||||  |||    | || || ||  ||||||| |||| ||||    | ||||   |
Db        370 AGGAAGTTGAGA---GCAGGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTT 426

Qy        468 CAGAAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGC 527
              | ||||||||||| |  || || | |||||||| |||||||  ||||  ||||| |||
Db        427 CGGAAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGCAGC 486

Qy        528 TTCAAGGGGCTGCAATTT 545
              || | |   ||| |  ||
Db        487 TTTAGGTTACTGGAGATT 504


RESULT 14
US-10-424-599-40645/c
; Sequence 40645, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
```

```
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 40645
;   LENGTH: 322
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(322)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1
US-10-424-599-40645

  Query Match             7.3%;  Score 125.6;  DB 10;  Length 322;
  Best Local Similarity   61.2%;  Pred. No. 6.3e-21;
  Matches  197;  Conservative    0;  Mismatches  125;  Indels    0;  Gaps    0;

Qy       1062 ATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTT 1121
             |||||  | ||  || |     |   |  | | ||  || | ||||| ||  ||||
Db        322 ATGGCANTAGAATAACTGTCACCCACCGCGCACAGTCAGCAGGAAGTTGACCACGACCCG 263

Qy       1122 GCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAG 1181
             |  ||| |||||||||      | |||  |  |  | | | ||||  ||||||||||||
Db        262 GACGATGTTGAAGATCCAGCCACGCTCCACCATCTGGCGCATGTTTCTAAAGATGAAAAA 203

Qy       1182 CAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCAT 1241
              | |||||||||| | ||  ||||  || | || ||  |||| ||      || ||||| ||
Db        202 TAGTTCATGCATNTNTGCCACTCTCTTATGAGGAAGAAAAGCGTCTGCCCACATGGTTAT 143

Qy       1242 ATCTCTTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCA 1301
              | |  ||||| ||||| ||||||       ||||| |  |||||| ||  |   |
Db        142 GTTTGNTGGGCTTGTGAGGCATTTCNCCAGTTTTATGGAAAAGAGCTGATCTTATCTCTT 83

Qy       1302 TCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGAC 1361
              |  | || ||| |||||||| || ||| |||| ||||| |||||  |||| ||| |  ||
Db         82 GCTTTATTTTGGCGTTGGAGGTTATTTATGATCAAACTTTGGAATTATGGTCTTCTTCAC 23

Qy       1362 TCAGCCACCATCAACAACTGCA 1383
              |    || || || || || |
Db         22 GCGCGTACTATGAATAANTGTA 1


RESULT 15
US-09-837-604B-6414
; Sequence 6414, Application US/09837604B
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  APPLICANT: De La Pena, Robert C.
;  APPLICANT:  Bougri, Olegs
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(51892)B
;  CURRENT APPLICATION NUMBER: US/09/837,604B
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: US 60/197,872
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 6414
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Oryza sativa nipponbare
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3431-030-P1-K2-G11
US-09-837-604B-6414

  Query Match             6.0%;  Score 104;  DB 5;  Length 453;
  Best Local Similarity   66.5%;  Pred. No. 1.6e-15;
  Matches  149;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy        424 CTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGTTA 483
              |||| |  ||  |||| ||||| ||  | |  ||||| ||  | || ||||| ||||| |
Db        173 CTGGCAATGTGCTTTTTAACTACAAATACTACAATAATACTATGCAAAAAACCGAAGTCA 232

Qy        484 GGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGGCTGCAAT 543
                || |||| |||||| |||||||||    | | |||||| ||||| ||||| ||    ||
Db        233 CTGAAGATTTTTCTTGCCCATTTTGCTTGGTACCATGTGGCAGCTTTAAGGGTCTAGGAT 292

Qy        544 TTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCGAAGAATACC 603
              |||  | ||  |||| |||||  | ||  | | ||||||   | | |  ||||| | ||
Db        293 GTCACCTAAACGCATCGCATGACCTTTTCCATTATGAGTTTTGGATATCTGAAGAGTGCC 352

Qy        604 AGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAG 647
              || | |||||||||  | | ||   | |||| |  | |   |||
Db        353 AGGCTGTTAATGTTAGTCTGAAGACTGATTCTTGGAGAACAGAG 396


Search completed: June 19, 2003, 23:39:28
Job time : 1299 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:22:23 ; Search time 72 Seconds
(without alignments)
823.563 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2378	100.0	445	21	AAB00060	VRN2 polypeptide.
2	2285	96.1	440	21	AAB00061	VRN2 polypeptide.
3	1923	80.9	498	21	AAG54168	Arabidopsis thalia
4	1796	75.5	367	21	AAG54169	Arabidopsis thalia
5	1635	68.8	339	21	AAG54170	Arabidopsis thalia
6	923	38.8	611	21	AAY53932	A MPC1 protein hav
7	906.5	38.1	604	21	AAY53933	An Os-MPC1 protein
8	549	23.1	107	21	AAB00062	VRN2 polypeptide g
9	510.5	21.5	813	21	AAB01674	FIS2 protein seque
10	451.5	19.0	186	21	AAB00065	At Hyp 2245035 (AT
11	368.5	15.5	108	21	AAB00064	AII93743 Expressed
12	327	13.8	63	21	AAG54171	Arabidopsis thalia
13	324	13.6	145	21	AAG61332	Arabidopsis thalia
14	324	13.6	174	21	AAG61331	Arabidopsis thalia
15	311	13.1	154	21	AAB00063	C72616 Expressed s
16	297	12.5	67	21	AAY53934	Fragment of a suga
17	225	9.5	388	22	AAU15978	Human novel secret
18	221	9.3	803	21	AAB00066	KIAA0160 polypepti
19	219	9.2	955	22	ABB71635	Drosophila melanog
20	153	6.4	289	22	AAU15958	Human novel secret
21	131.5	5.5	1245	21	AAB18244	Plasmodium falcipa
22	130.5	5.5	710	22	AAB30814	Amino acid sequenc
23	129	5.4	22	21	AAB00110	Zinc finger motif
24	128.5	5.4	1817	21	AAB18301	Plasmodium falcipa
25	127.5	5.4	700	21	AAY58359	Yeast G protein-co
26	127.5	5.4	847	21	AAY58360	Yeast G protein-co
27	127.5	5.4	953	21	AAY58356	Yeast G protein-co
28	127.5	5.4	953	21	AAY58357	Yeast G protein-co
29	127.5	5.4	961	21	AAY58355	Yeast G protein-co
30	125.5	5.3	1501	23	AAU76762	Plasmodium falcipa
31	124.5	5.2	686	21	AAB18293	Plasmodium falcipa
32	123.5	5.2	2010	21	AAB18218	Plasmodium falcipa
33	122	5.1	190	14	AAR41998	Staphylococcus ent
34	122	5.1	190	14	AAR42009	Staphylococcus ent
35	122	5.1	256	13	AAR24796	Sequence of fibron
36	122	5.1	1714	21	AAB18275	Plasmodium falcipa
37	121	5.1	190	14	AAR42004	Staphylococcus ent
38	121	5.1	590	12	AAR11019	Apo-lipoprotein E
39	120.5	5.1	1436	21	AAB18199	Plasmodium falcipa
40	120	5.0	190	14	AAR42005	Staphylococcus ent
41	120	5.0	2380	21	AAB18315	Plasmodium falcipa
42	119.5	5.0	190	14	AAR42011	Staphylococcus ent
43	119	5.0	190	14	AAR41989	Staphylococcus ent
44	119	5.0	190	14	AAR41990	Staphylococcus ent
45	119	5.0	190	14	AAR41991	Staphylococcus ent

ALIGNMENTS

```
RESULT 1
AAB00060
ID   AAB00060 standard; Protein; 445 AA.
XX
AC   AAB00060;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 polypeptide.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation.
XX
OS   Arabidopsis thaliana var 'Landsberg erecta'.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   57..61
FT                   /note= "Putative NLS"
FT   Region          90..111
FT                   /label= Zinc finger motif
FT   Modified-site   158..159
FT                   /note= "Putative NLS "
FT   Modified-site   170..174
FT                   /note= "Putative NLS"
FT   Domain          281..316
FT                   /note= "Putative acidic activation domain"
FT   Misc-difference 323
FT                   /note= "mutation in mutant gene vin2-1 causes a change
FT                           from a Trp codon to a STOP codon resulting in a
FT                           truncated mutant protein"
XX
PN   WO200044918-A1.
```

```
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   N-PSDB; AAA47751.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 49; Fig 6; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence   445 AA;

  Query Match             100.0%;  Score 2378;  DB 21;  Length 445;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-234;
  Matches  445;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420

Qy        421 DHPNDINNKNNVDNKDNNSRDKVIK 445
              |||||||||||||||||||||||||
Db        421 DHPNDINNKNNVDNKDNNSRDKVIK 445


RESULT 2
AAB00061
ID   AAB00061 standard; Protein; 440 AA.
XX
AC   AAB00061;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 polypeptide.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation.
XX
OS   Arabidopsis thaliana (Columbia).
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   N-PSDB; AAA47753.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 47; Page 71-72; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence   440 AA;

  Query Match             96.1%;  Score 2285;  DB 21;  Length 440;
  Best Local Similarity   96.8%;  Pred. No. 7.8e-225;
  Matches  428;  Conservative    3;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240
              |||| ||||||| |||||||||||||||||||||||||||||||||||||||||||||:|
Db        181 SPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||
Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEVFSRFYEKELHCYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420
              ||||||||||||||||||||||||||||||||||:|    | ||||||| ||||||||||
Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNTS--VTNNNNNSVDHPSDSNTNNNNIV 418
```

```
Qy        421 DHPNDINNKNNVDNKDNNSRDK 442
              |||||| ||||||||||||||||
Db        419 DHPNDIKNKNNVDNKDNNSRDK 440


RESULT 3
AAG54168
ID   AAG54168 standard; Protein; 498 AA.
XX
AC   AAG54168;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 69039.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
```

```
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match            80.9%;  Score 1923;  DB 21;  Length 498;
  Best Local Similarity  89.4%;  Pred. No. 1.1e-187;
  Matches  370;  Conservative    3;  Mismatches    7;  Indels   34;  Gaps    3;

Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 167

Qy        103 LQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQ 162
              |||||||||||||||||    | :|| ::   |                  ||||||||
Db        168 LQFHLNSSHDLFEFEFK----YISVNPTIMFCS------------------SKPRKRRQ 204

Qy        163 RGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPP 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 RGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPP 264

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEAR-----------SHLLLQKRQFYH 271
              ||||||||||||||||||||||||||||||||||||||           ||||||||||||
Db        265 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSTAILTEKISSSHLLLQKRQFYH 324

Qy        272 SHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQR 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 SHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQR 384

Qy        332 VIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENC 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 VIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENW 444

Qy        392 RNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        445 RNSSDTTTTNNNNSVDRPSDSNTNNNNIVSHPNDINNKNNVDNKDNNSRDKVIK 498
```

RESULT 4
AAG54169
ID AAG54169 standard; Protein; 367 AA.
XX
AC AAG54169;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69040.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             75.5%;  Score 1796;  DB 21;  Length 367;
  Best Local Similarity   88.7%;  Pred. No. 6.8e-175;
  Matches  346;  Conservative    3;  Mismatches    7;  Indels   34;  Gaps    3;

Qy         67 MVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQT 126
              |||||||||||||||||||||||||||||||||||||||||||||||||||||    | :
Db          1 MVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFK----YIS 56

Qy        127 VNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
              || ::   |                  ||||||||||||||||||||||||||||||||
Db         57 VNPTIMFCS------------------SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 97

Qy        187 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 157

Qy        247 KTRKLSAERSEAR-----------SHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDD 295
              ||||||||||||            ||||||||||||||||||||||||||||||||||
Db        158 KTRKLSAERSEARSTAILTEKISSSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDD 217

Qy        296 DVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHR 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 DVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHR 277

Qy        356 YSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTN 415
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        278 YSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENWRNSSDTTTTNNNNSVDRPSDSNTN 337

Qy        416 NNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
              ||||| ||||||||||||||||||||||||
Db        338 NNNIVSHPNDINNKNNVDNKDNNSRDKVIK 367


RESULT 5
AAG54170
ID   AAG54170 standard; Protein; 339 AA.
XX
AC   AAG54170;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 69041.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.8%; Score 1635; DB 21; Length 339;
Best Local Similarity 87.8%; Pred. No. 1.8e-158;
Matches 318; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```
Qy         95 MLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLC 154
              |||||||||||||||||||||||||    | :|| ::   |
Db          1 MLCGSFKGLQFHLNSSHDLFEFEFK----YISVNPTIMFCS------------------- 37

Qy        155 SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQF 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQF 97

Qy        215 EMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEAR----------SHLL 263
              |||||||||||||||||||||||||||||||||||||||||||||          ||||
Db         98 EMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSTAILTEKISSSHLL 157

Qy        264 LQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLW 323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 LQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLW 217

Qy        324 NSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINN 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 NSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINN 277

Qy        384 CNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKV 443
              ||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        278 CNTILENWRNSSDTTTTNNNNSVDRPSDSNTNNNNIVSHPNDINNKNNVDNKDNNSRDKV 337

Qy        444 IK 445
              ||
Db        338 IK 339
```

RESULT 6
AAY53932
ID AAY53932 standard; Protein; 611 AA.
XX
AC AAY53932;
XX
DT 13-MAR-2000 (first entry)
XX
DE A MPC1 protein having flowering regulating activity.
XX
KW Flowering regulating activity; MPC1; flowering; germination;
KW super early flowering mutation; altered flowering time;
KW flowering regulating gene; food crop; vegetable; flowering inhibition;
KW productivity.
XX
OS Arabidopsis thaliana.
XX

```
FH   Key             Location/Qualifiers
FT   Domain          306..327
FT                   /note= "zinc finger domain"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;    99EP-0305077.
XX
PR   26-JUN-1998;    98JP-0180065.
PR   24-JUN-1999;    99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   N-PSDB; AAZ36947, AAZ36948.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 2; Page 12-15; 53pp; English.
XX
CC   The present sequence represents a protein having a flowering regulating
CC   activity, which is designated MPC1. The genomic sequence is given
CC   in AAZ36948. A rice MPC1 is also disclosed in the specification. The
CC   rice and Arabidopsis cDNAs show significant homology with each other.
CC   A naturally occurring mutation of the MPC1 gene eliminates normal
CC   flowering regulating ability of plants, and leads to flowering
CC   immediately after germination (super early flowering mutation), and
CC   leads to a truncated protein comprising amino acids 1-540 of the present
CC   sequence. The MPC1 polynucleotide sequence can be used to produce plants
CC   with altered flowering times in comparison with wild type plants, by
CC   enhancing or inhibiting the expression of the flowering regulating gene.
CC   Antisense polynucleotides can be used to reproduce the effects of the
CC   mutated MPC1 gene. This alteration can be used to increase the yield
CC   of food crops. Flowering inhibition of vegetables increases their
CC   productivity.
XX
SQ   Sequence   611 AA;

  Query Match             38.8%;  Score 923;  DB 21;  Length 611;
  Best Local Similarity   34.8%;  Pred. No. 3.5e-85;
  Matches  216;  Conservative   52;  Mismatches  108;  Indels  244;  Gaps    9;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              || :: | : | || |: :|:|  ||||| ||||   |:: || || |||:||| || ||
Db          1 MCHEDSRLRISEEEEIAAEESLAAYCKPVELYNIIQRRAIRNPLFLQRCLHYKIEAKHKR 60

Qy         61 ------------------------------------------------------------ 60

Db         61 RIQMTVFLSGAIDAGVQTQKLFPLYILLARLVSPKPVAEYSAVYRFSRACILTGGLGVDG 120

Qy         61 ------------------------------------------------------------ 60

Db        121 VSQAQANFLLPDMNRLALEAKSGSLAILFISFAGAQNSQFGIDSGKIHSGNIGGHCLWSK 180

Qy         61 -----------------------------------KSRS------------------- 64
                                                  || |
Db        181 IPLQSLYASWQKSPNMDLGQRVDTVSLVEMQPCFIKLKSMSEEKCVSIQVPSNPLTSSSP 240

Qy         65 ---------------------------------------TGMVVFNYKDCNNTLQKTEV 84
                                                     || |||||:  || ||||||
Db        241 QQVQVTISAEEVGSTEKSPYSSFSYNDISSSSLLQIIRLRTGNVVFNYRYYNNKLQKTEV 300

Qy         85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSD 144
               || ||||| : | |||||::|| |:|||  ||| : ||:| ||||:|  : | :    |
Db        301 TEDFSCPFCLVKCASFKGLRYHLPSTHDLLNFEFWVTEEFQAVNVSLKTETMISKVNEDD 360

Qy        145 -DDKFEPFSLCSKP-RKRRQRGGRNNTRRLKVCFLPLDSPSLTNG------TENGITLLN 196
              | | : |   ||  |:|||:    ::|:          | |  |      |::  :: :

Db        361 VDPKQQTFFFSSKKFRRRRQKSQVRSSRQ---------GPHLGLGCEVLDKTDDAHSVRS 411

Qy        197 DGNR---GLGYPEATELAGQFEMTSNIP----PAIAHSSLDAGAKVILTSEAVVPATKTR 249
              : :|   | :|    |  |  |    :|    ||   |  |      :    ::   |||
Db        412 EKSRIPPGKHY----ERIGGAESGQRVPPGTSPADVQSCGDPDYVQSIAGSTMLQFAKTR 467

Qy        250 KLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDF 309
              |:| |||: |:  |||||||:|||| ||||||||:||||||||||||||||||:|||||
Db        468 KISIERSDLRNRSLLQKRQFFHSHRAQPMALEQVLSDRDSEDEVDDDVADFEDRRMLDDF 527

Qy        310 VDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIK 369
              ||| ||||| ||:||||||||||:||||| |||||||| :   : |   | ||||:|::|
Db        528 VDVTKDEKQMMHMWNSFVRKQRVLADGHIPWACEAFSRLHGPIMVRTPHLIWCWRVFMVK 587

Qy        370 LWNHGLVDSATINNCNTILE 389
              ||||||:|: |:||||| ||
Db        588 LWNHGLLDARTMNNCNTFLE 607


RESULT 7
AAY53933
ID   AAY53933 standard; Protein; 604 AA.
XX
AC   AAY53933;
XX
DT   13-MAR-2000  (first entry)
XX
DE   An Os-MPC1 protein having flowering regulating activity.
XX
KW   Flowering regulating activity; MPC1; flowering; germination;
KW   super early flowering mutation; altered flowering time;
KW   flowering regulating gene; food crop; vegetable; flowering inhibition;
KW   productivity.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   Domain          310..335
FT                   /note= "zinc finger domain"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;    99EP-0305077.
XX
PR   26-JUN-1998;    98JP-0180065.
PR   24-JUN-1999;    99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   N-PSDB; AAZ36953.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 3; Page 36-39; 53pp; English.
XX
CC   The present sequence represents a protein having a flowering regulating
CC   activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
CC   disclosed in the specification. The rice and Arabidopsis cDNAs show
CC   significant homology with each other. A naturally occurring mutation
CC   of the MPC1 gene eliminates normal flowering regulating ability of
CC   plants, and leads to flowering immediately after germination (super
CC   early flowering mutation). The MPC1 polynucleotide sequence can be
CC   used to produce plants with altered flowering times in comparison
CC   with wild type plants, by enhancing or inhibiting the expression of
CC   the flowering regulating gene. Antisense polynucleotides can be used
CC   to reproduce the effects of the mutated MPC1 gene. This alteration
CC   can be used to increase the yield of food crops. Flowering inhibition
```

```
CC   of vegetables increases their productivity.
XX
SQ   Sequence   604 AA;

  Query Match             38.1%;  Score 906.5;  DB 21;  Length 604;
  Best Local Similarity   33.3%;  Pred. No. 1.7e-83;
  Matches  209;  Conservative   59;  Mismatches  109;  Indels  251;  Gaps    7;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||   ||: ||:| :: :|:  :|||| ||||   ||: ||:|| ||| ||| |:||:
Db          1 MCRHQPRARLSPDEQLAAEESFALYCKPVELYNIIQRRSIKNPAFLQRCLLYKIHARRKK 60

Qy         61 KSRST---------------------------------------------------- 65
              :|  |
Db         61 RSLITISLSGGTNKELRAQNIFPLYVLLARPTNNVSLEGHSPIYRFSRACLLTSFHEFGN 120

Qy         66 ---------------------------------------------------------- 65

Db        121 KDYTEATFVIPDVKNLATSRACSLNIILISCGRAEQTFDDNNCSGNHVEGSTLQKLEGKC 180

Qy         66 ---------------------------------------------------------- 65

Db        181 FWGKIPIDLLASSLGNCVSLSLGHTVEMSSTVEMTPSFLEPKFLEDDSCLTFCSQKVDAT 240

Qy         66 ----------------------------------------------GMVVFNYKDCNNTLQ 80
                                                           | |:|||| |||:|
Db        241 GSFQLQVSISAQEAGAKDMSESPYSVYSYNDVPPSSLTHIIRLRSGNVLFNYKYYNNTMQ 300

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFE- 139
              |||| || ||||| : ||||||| |||:||||| :|| : || | ||||:| :|:  |
Db        301 KTEVTEDFSCPFCLVPCGSFKGLGCHLNASHDLFHYEFWISEECQAVNVSLKTDSWRTEL 360

Qy        140 -EEGSDDDKFEPFSLCSKPRKR-RQRGGRNNTRRLKVCFLPLDSP-SLTNGTENGITLLN 196
               ||  | : : ||  |: :|| |     :  | :    :   ||     |:|:
Db        361 LAEGV-DPRHQTFSYRSRFKKRKRVEISSDKIRHVHPHIVDSGSPEDAQAGSEDDYVQRE 419

Qy        197 DGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKV---ILTSEAVVPATKTRKLSA 253
              :|:                          ::||:|:|    :    |::  |:   ||||||
Db        420 NGS----------------------SVAHASVDPANSLHGSNLSAPTVLQFGKTRKLSV 456

Qy        254 ERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVN 313
              ||:: |:  |||||||:|||| ||||  :| ||||||||||||:||||||:||||||||
Db        457 ERADPRNRQLLQKRQFFHSHRAQPMAWSKVFSDRDSEDEVDDDIADFEDRRMLDDFVDVT 516

Qy        314 KDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNH 373
              ||||  ||:|||||||||:||||| |||||||:|: :|| :  :| |||| |::|||||
Db        517 KDEKLIMHMWNSFVRKQRVLADGHIPWACEAFSQFHGQELVQNPALLWCWRFFMVKLWNH 576

Qy        374 GLVDSATINNCNTILENCRNSSDTTTTN 401
               |:|:  :| ||||||   | |     |
Db        577 SLLDARAMNACNTILEGYLNGSSDPKKN 604


RESULT 8
AAB00062
ID   AAB00062 standard; Protein; 107 AA.
XX
AC   AAB00062;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 polypeptide generated from aberrantly spliced VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation.
XX
OS   Arabidopsis thaliana (Columbia).
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   N-PSDB; AAA47755.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 47; Page 75; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant. This polypeptide resulted from an aberrant splice in the
CC   VRN2 gene.
XX
SQ   Sequence   107 AA;

  Query Match             23.1%;  Score 549;  DB 21;  Length 107;
  Best Local Similarity   99.0%;  Pred. No. 4.9e-48;
  Matches  100;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101
              |||||||||||||||||||:|||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFK 101


RESULT 9
AAB01674
ID   AAB01674 standard; protein; 813 AA.
XX
AC   AAB01674;
XX
DT   01-SEP-2000  (first entry)
XX
DE   FIS2 protein sequence.
XX
KW   Seed; development; FIS; endosperm; autonomous embryogenesis;
KW   transgenic plant; seedless fruit; parthenocarpic; citrus fruit;
KW   stone fruit.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..120
FT                   /note= "The first 120 residues are illegible in the
FT                   specification"
XX
PN   WO200016609-A1.
XX
PD   30-MAR-2000.
XX
PF   21-SEP-1999;   99WO-AU00805.
XX
PR   21-SEP-1998;   98US-0101184.
PR   22-SEP-1998;   98AU-0006061.
PR   22-SEP-1998;   98AU-0006062.
PR   22-SEP-1998;   98AU-0006063.
PR   01-JUL-1999;   99AU-0001345.
```

PR 01-JUL-1999; 99AU-0001346.
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
PI Bilodeau P, Chaudhury AM, Dennis ES, Koltunow AMG, Luo M;
PI Peacock WJ;
XX
DR WPI; 2000-283392/24.
XX
PT Induction of seed development in plants in the absence of fertilization
PT by inhibiting or preventing the expression of a negative regulator of
PT seed formation for production of seedless or soft-seeded fruit -
XX
PS Claim 2; Fig 14; 207pp; English.
XX
CC The present invention relates to a method of inducing the development of
CC seeds in a plant, comprising inhibiting, interrupting or reducing the
CC expression of a negative regulator of seed formation in one or more
CC female reproductive cells, tissues, or organs of the plant or a
CC progenitor cell, tissue or organ. The negative regulator is a
CC polypeptide. The present sequence represents the FIS2 protein sequence.
CC The FIS family of genes are known to be capable of regulating autonomous
CC endosperm development and/or autonomous embryogenesis. In the invention
CC the reduced expression of the negative regulator is achieved by the
CC introduction of a transgene which comprises a FIS genetic sequence, which
CC may inhibit FIS activity. Plants produced using the method of the
CC invention produce parthenocarpic fruit or soft-seeded fruit, where the
CC fruit are made parthenocarpic or have soft seed by a process comprising
CC expressing the introduced nucleic acid molecule in a tissue or organ of
CC the fruit. The plant produces seed independent of fertilization. The
CC isolated FIS nucleic acid molecules are used in the production of an
CC antisense molecule, a ribozyme, a co-suppression molecule, a
CC gene-targeting molecule, a gene-silencing molecule and a
CC dominant-negative sense molecule where the member is used for the
CC production of a transformed plant. The transformed plant is apomictic or
CC produces soft-seeded or parthenocarpic fruit. Production of soft-seeded
CC fruit has large economic value, since it makes the fruit more desirable
CC to customers. Examples include stone fruits such as apricots and peaches,
CC citrus fruits such as oranges, lemons, grapefruit and mandarins and other
CC fruits such as grapes, apples, melons, pears and berries. The plants
CC which undergo autonomous seed formation do not require fertilization to
CC reproduce, and may express desirable characteristics stably between
CC generations. Antibodies produced to the FIS polypeptides can be used to
CC detect the peptides of the invention and can be used in an enzyme linked
CC immunosorbant assay (ELISA), radioimmunoassay or histochemical tests.
XX
SQ Sequence 813 AA;

Query Match 21.5%; Score 510.5; DB 21; Length 813;
Best Local Similarity 22.3%; Pred. No. 9e-43;
Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;

```
Qy          81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVSVKLNSFIF 138
               | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|||| ::  |
Db         125 KAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDVSVKSDTIKF 184

Qy         139 E--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195
                 ::   : :  | : ||| | :|||| || ::| |  : ||   |  ||||  | :
Db         185 GVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPRGTENDSTHV 244

Qy         196 NDGN---------------------------------------------------------- 199
               || |
Db         245 NDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAHSSAEKNEST 304

Qy         200 ---------------------------------------------------------- 199

Db         305 HVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDEDVSFPPRTRS 364

Qy         200 ------------------------RG---------------------------------- 201
                                       ||
Db         365 SKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSEPKVLHVNDE 424

Qy         202 --LGYPEATELAGQFEMTSNIPPAIAHSS------------------------------ 228
```

```
                 |||  |    :: :   |||| ||
Db         425 NVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSKKNKSTRKNV 484

Qy         229 ---------------------------------------------------------- 228

Db         485 DNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSSKKNKSTRKN 544

Qy         229 -------------------LDAGAKVILTSEAVVP------------------------ 244
                               |      |  ||  ||
Db         545 DDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSSKKNKSTHKK 604

Qy         245 -----------------------------------ATKTRKLSAERSEARSHLLLQKR 267
                                                  :: ::| ||| ||:    |: |
Db         605 DDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEAKRLERLKGR 664

Qy         268 QFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
               |||||  :|||  ||||||: |||:| ||   |  :|  |:  | |:|:||::|:|||. ||
Db         665 QFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKRYMYLWNIFV 724

Qy         328 RKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTI 387
               ||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: : | : | ||
Db         725 RKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICAKTFHKCTTI 784

Qy         388 LENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
               |           :|:: :   | |  | ||
Db         785 L----------LSNSDEAGQFTSGSAANANN 805
```

RESULT 10
AAB00065
ID AAB00065 standard; Protein; 186 AA.
XX
AC AAB00065;
XX
DT 16-NOV-2000 (first entry)
XX
DE At Hyp 2245035 (ATFCA7_4) (modified).
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time;
KW leaf size; leaf shape; shade avoidance response; reproduction;
KW breeding; pollination; cultivation.
XX
OS Unspecified.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
DR N-PSDB; AAA47758.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 77; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA was generated from an aberrantly spliced VRN2
CC gene.

```
XX
SQ   Sequence   186 AA;

  Query Match           19.0%;  Score 451.5;  DB 21;  Length 186;
  Best Local Similarity   50.8%;  Pred. No. 1.1e-37;
  Matches   99;  Conservative   32;  Mismatches   43;  Indels   21;  Gaps    7;

Qy        237 LTSEAVVPATKTRKLSAERSEARSHLL-LQKRQFYHSHRVQPMALEQVMSDRDSEDEVD- 294
              ||:|| ||        |:||:| || | | |||||||   ||::||||||||||||::|| 
Db          6 LTTEAKVP--------AKRSKATSHYLPLHKRQFYHSRTGQPLSLEQVMSDRDSENDVDK 57

Qy        295 -DDVADFEDRQMLDDFVDVNK-DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKE 352
               || |  |: |||:  :| |:   ::|: ||||||::||::|| || ||||||||  : :| 
Db         58 NDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAHIPWACEAFSRLHLQE 117

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDS 412
              |    ||  ||| |:|| |::||:|  |:| ||||: :  | | |    |||:
Db        118 LRSNLSLDLCWRQFMIKQWDYGLLDRVTMNKCNTIIYH--NISTTNDDINNNNT-----R 170

Qy        413 NTNNNNIVDHPNDIN 427
              |:| ::||  :|||
Db        171 TTDNMDVVD--DDIN 183


RESULT 11
AAB00064
ID   AAB00064 standard; Protein; 108 AA.
XX
AC   AAB00064;
XX
DT   16-NOV-2000  (first entry)
XX
DE   AII93743 Expressed sequence tag (modified).
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation.
XX
OS   Unspecified.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   N-PSDB; AAA47757.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Disclosure; Page 76; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant. This cDNA was generated from an aberrantly spliced VRN2
CC   gene.
XX
SQ   Sequence   108 AA;

  Query Match           15.5%;  Score 368.5;  DB 21;  Length 108;
  Best Local Similarity   60.8%;  Pred. No. 1.5e-29;
  Matches   76;  Conservative    9;  Mismatches   21;  Indels   19;  Gaps    2;

Qy        204 YPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLL 263
              ||:|     |    :|: |                  |::   |||||| |||: |:  |
Db          3 YPDAE--CAQLVPGNNLAP----------------PAMLQFAKTRKLSIERSDMRNRTL 43

Qy        264 LQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLW 323
              | ||||:|||| |||| ||||||||||||||||||||||:|||||||| ||||| ||||
Db         44 LHKRQFFHSHRAQPMAAEQVMSDRDSEDEVDDDVADFEDRRMLDDFVDVTKDEKQMMHLW 103

Qy        324 NSFVR 328
              |||||
Db        104 NSFVR 108


RESULT 12
AAG54171
ID   AAG54171 standard; Protein; 63 AA.
XX
AC   AAG54171;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 69043.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
```

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```
  Query Match             13.8%;  Score 327;  DB 21;  Length 63;
  Best Local Similarity   96.8%;  Pred. No. 1.2e-25;
  Matches   61;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        383 NCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDK 442
              |||||||| |||||||||||||||||||||||||||| |||||||||||||||||||||
Db          1 NCNTILENWRNSSDTTTTNNNNSVDRPSDSNTNNNNIVSHPNDINNKNNVDNKDNNSRDK 60

Qy        443 VIK 445
              |||
Db         61 VIK 63


RESULT 13
AAG61332
ID   AAG61332 standard; Protein; 145 AA.
XX
AC   AAG61332;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 79541.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
```

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 324; DB 21; Length 145;
Best Local Similarity 98.3%; Pred. No. 8.3e-25;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db         81 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFK 139
```

RESULT 14
AAG61331
ID AAG61331 standard; Protein; 174 AA.
XX
AC AAG61331;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79540.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 324; DB 21; Length 174;
Best Local Similarity 98.3%; Pred. No. 1.1e-24;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101
||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db 110 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFK 168

RESULT 15

AAB00063
ID AAB00063 standard; Protein; 154 AA.
XX
AC AAB00063;
XX
DT 16-NOV-2000 (first entry)
XX
DE C72616 Expressed sequence tag (modified).
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time;
KW leaf size; leaf shape; shade avoidance response; reproduction;
KW breeding; pollination; cultivation.
XX
OS Unspecified.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
DR N-PSDB; AAA47756.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 76; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA was generated from an aberrantly spliced VRN2
CC gene.
XX
SQ Sequence 154 AA;

Query Match 13.1%; Score 311; DB 21; Length 154;
Best Local Similarity 62.1%; Pred. No. 2e-23;
Matches 64; Conservative 14; Mismatches 21; Indels 4; Gaps 2;

```
Qy        223 AIAHSSLDAGAKV---ILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMA 279
              ::||:|:|    :    |::  |:  |||||| ||:: |:  |||||||:|||| ||||
Db         53 SVAHASVDPANSLHGSNLSAPTVLQFGKTRKLSVERADPRNRQLLQKRQFFHSHRAQPMA 112

Qy        280 LEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHL 322
              |  | ||||||||||||:||||||||||||||| |||   ||:
Db        113 LGAVFSDRDSEDEVDDDIADFEDRQMLDDFVDVTKDE-LIMHM 154
```

Search completed: June 13, 2003, 15:35:41
Job time : 74 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:37:19 ; Search time 26 Seconds
(without alignments)
503.585 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	126.5	5.3	1584	4	US-09-457-040B-27	Sequence 27, Appl
2	123	5.2	110	2	US-08-569-166-34	Sequence 34, Appl
3	112	4.7	361	1	US-08-415-751-3	Sequence 3, Appli
4	105	4.4	732	4	US-08-914-999-8	Sequence 8, Appli
5	103	4.3	432	2	US-08-700-152A-4	Sequence 4, Appli
6	103	4.3	1073	4	US-09-541-782-6	Sequence 6, Appli
7	103	4.3	1073	4	US-09-723-820-6	Sequence 6, Appli
8	98.5	4.1	1388	4	US-09-572-191-2	Sequence 2, Appli
9	98.5	4.1	1388	4	US-09-723-262-2	Sequence 2, Appli
10	98.5	4.1	1388	4	US-09-723-219-2	Sequence 2, Appli
11	98	4.1	652	4	US-08-559-896B-2	Sequence 2, Appli
12	94	4.0	2391	2	US-08-446-855A-2	Sequence 2, Appli
13	94	4.0	2391	4	US-09-150-741-2	Sequence 2, Appli
14	92	3.9	286	4	US-09-384-162-6	Sequence 6, Appli
15	91.5	3.8	689	4	US-09-177-249-2	Sequence 2, Appli
16	91.5	3.8	689	4	US-09-061-769A-2	Sequence 2, Appli
17	91.5	3.8	832	2	US-08-677-734A-12	Sequence 12, Appl
18	91.5	3.8	832	4	US-09-097-053-12	Sequence 12, Appl
19	91.5	3.8	1786	4	US-08-973-462-8	Sequence 8, Appli
20	91	3.8	1051	4	US-08-965-762-14	Sequence 14, Appl
21	91	3.8	1051	4	US-09-911-927-14	Sequence 14, Appl
22	91	3.8	1051	4	US-09-911-882-14	Sequence 14, Appl
23	91	3.8	1099	4	US-09-442-100-2	Sequence 2, Appli
24	90.5	3.8	428	3	US-09-118-319-5	Sequence 5, Appli
25	90.5	3.8	581	3	US-08-619-812-6	Sequence 6, Appli
26	89.5	3.8	826	1	US-07-638-431-2	Sequence 2, Appli
27	89.5	3.8	826	5	PCT-US92-00018-2	Sequence 2, Appli
28	89	3.7	749	2	US-08-568-459A-6	Sequence 6, Appli
29	89	3.7	749	2	US-08-487-826B-6	Sequence 6, Appli
30	89	3.7	749	4	US-09-210-288-6	Sequence 6, Appli
31	88.5	3.7	432	3	US-09-118-319-2	Sequence 2, Appli
32	88.5	3.7	507	4	US-09-134-001C-3978	Sequence 3978, Ap
33	88.5	3.7	596	4	US-09-752-165-2	Sequence 2, Appli
34	88.5	3.7	911	3	US-09-074-579-4	Sequence 4, Appli
35	88.5	3.7	911	4	US-09-388-774-4	Sequence 4, Appli
36	88	3.7	930	4	US-09-134-001C-5314	Sequence 5314, Ap
37	87.5	3.7	392	3	US-08-523-373-6	Sequence 6, Appli
38	87.5	3.7	1081	2	US-08-843-530B-18	Sequence 18, Appl
39	87.5	3.7	1117	2	US-08-843-530B-33	Sequence 33, Appl
40	87	3.7	620	1	US-08-484-105-4	Sequence 4, Appli
41	87	3.7	620	1	US-08-484-106-4	Sequence 4, Appli
42	87	3.7	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
43	86.5	3.6	770	2	US-08-209-521-13	Sequence 13, Appl
44	86.5	3.6	770	4	US-08-961-810-123	Sequence 123, App
45	86.5	3.6	770	4	US-08-352-902D-123	Sequence 123, App

ALIGNMENTS

RESULT 1
US-09-457-040B-27
; Sequence 27, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: DICDI - Dictyostelium Discoideum
US-09-457-040B-27

Query Match 5.3%; Score 126.5; DB 4; Length 1584;
Best Local Similarity 22.2%; Pred. No. 0.00056;
Matches 77; Conservative 37; Mismatches 100; Indels 133; Gaps 16;

```
Qy        149 EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEAT 208
              | | :|    |:    |::   | :  |||||||  | |   |:        |  |
Db        210 EEFEVCRWSEKKNYH-GKHVVVRNRTAFLPLDSPKDTIG---GVRATQPFGEGFCY---- 261

Qy        209 ELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQ 268
                    ||        :    || |   |  :    |                |:    |
Db        262 -----FE--------VIIDQLDKGQLSIGLANLEYPTF------------YHVGWMPRS 295

Qy        269 F-YHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
              : ||                        :||   |  |:       ||: |      : |
Db        296 YGYH----------------------NDDGRKFRWREE----PGVNEGES-----YGSSY 324

Qy        328 RKQRVIADGHISWACE-------------AFSRFYEKELHRYSSLFWCWRLF-------- 366
              :|  :|  |.:|:                |||  |         :|:    |
Db        325 KKGDIIGCG-LSFTSREIFFTKNGMYLGTAFSNVY--------GVFYPSVAFNEPGISIT 375

Qy        367 ------------LIKLWNHGLVDSATI----------------NNCNTILENCRNSSDTT 398
                          : | |   |:| :|                || | |: | : :: ||
Db        376 GVFGPPFKFSQVTLMLKN---VNSTSILVPNGNNNNNSNNNNNNNNNNIIGNGKITTTTT 432

Qy        399 TT------NNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNS 439
              |:      |||  :   :::| ||||  :: |: || || :| :|||
Db        433 TSTSPSSINNNEDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNS 479
```

RESULT 2
US-08-569-166-34

```
; Sequence 34, Application US/08569166
; Patent No. 5830722
;  GENERAL INFORMATION:
;    APPLICANT:  NICOLAS, LUC
;    APPLICANT:  CHARLES, JEAN-FRANCOIS
;    APPLICANT:  DELECLUSE, ARMELLE
;    APPLICANT:  BARLOY, FREDERIQUE
;    TITLE OF INVENTION:  CLOSTRIDIUM BIFERMENTANS DNA FRAGMENT
;    TITLE OF INVENTION:  BEARING GENES CODING FOR PROTEINS LINKED TO AN
;    TITLE OF INVENTION:  INSECTICIDAL ACTIVITY
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/569,166
;      FILING DATE:  05-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR94/00768
;      FILING DATE:  24-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 93/07795
;      FILING DATE:  25-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  660-106-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  110 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-569-166-34

  Query Match             5.2%;  Score 123;  DB 2;  Length 110;
  Best Local Similarity   38.7%;  Pred. No. 1.8e-05;
  Matches   24;  Conservative   16;  Mismatches   18;  Indels    4;  Gaps    1;

Qy        380 TINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNS 439
              | || |    |  |::: |  ||||:    :: ||||||: :: |: || || ::::||:
Db         10 TNNNTNNNNNNINNNNNNTNNNNNA----NNQNTNNNNMGNNSNNNNNPNNNNHQNNNN 65

Qy      *  440 RD 441
              :
Db         66 NN 67


RESULT 3
US-08-415-751-3
; Sequence 3, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  361 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
;    FEATURE:
;      NAME/KEY:  Positions coded by nonsense codons are
;      NAME/KEY:  identified as Xaa.
US-08-415-751-3

  Query Match             4.7%;  Score 112;  DB 1;  Length 361;
  Best Local Similarity   38.3%;  Pred. No. 0.0019;
  Matches   23;  Conservative   12;  Mismatches   25;  Indels    0;  Gaps    0;

Qy        382 NNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRD 441
              || |    |  |::  |  ||||   | : :: |||||  :  : |||:| :| :|:: |
Db        115 NNHNNSXSNHYNNNSKTNNNNNNKGSRXATNSHNNNNIKANSYNNNNKSNNNNNNNSAND 174


RESULT 4
US-08-914-999-8
; Sequence 8, Application US/08914999
; Patent No. 6346406
;  GENERAL INFORMATION:
;    APPLICANT:  Ryazanov, Alexey G.
;    APPLICANT:  Hait, William N.
;    APPLICANT:  Pavur, Karen S.
;    TITLE OF INVENTION:  ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
;    TITLE OF INVENTION:  AND METHODS OF USE THEREFOR
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David A. Jackson, Esq.
;      STREET:  411 Hackensack Ave, Continental Plaza, 4th
;      STREET:  Floor
;      CITY:  Hackensack
```

```
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/914,999
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  601-1-078
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201-343-1684
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  732 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Dictyostelium discoideum
US-08-914-999-8

  Query Match             4.4%;  Score 105;  DB 4;  Length 732;
  Best Local Similarity   20.1%;  Pred. No. 0.033;
  Matches   76;  Conservative   61;  Mismatches  167;  Indels   74;  Gaps   17;

Qy        117 EFKLFE-EYQTVNVSVKLNSFIFEEEG---SDDDKFEPFSLCSKPRKRRQRGGRNNTRRL 172
              | |:|   | |    :::   :    |   : | ||| |   :      : : |     |:
Db         34 EGKIFRFPYGTSQTDLQIGKMLPSGSGGGATADSKFEKFKARNTLADIQYKVGDTLYVRV 93

Qy        173 KV------CFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNI----PP 222
              |          ||  : :  :|:|  |    |       | |    |:  |: :    |
Db         94 KKSKPTNDSLLPTLNIAFLDGSERAIKWEYD-------PYTT--TAQWTCTATLVKVEPV 144

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQF--------YHSHR 274
              |  :       :  |::        |:  : :   |:    | |          |:| :
Db        145 PFAEGAFRKAYHTLDLSKSGASGRYVSKIGKKPTPRPSYFEDVKMQMIAKKWADKYNSFK 204

Qy        275 VQPMALEQVMS------DRDSEDEV--DDDVADFEDRQMLDDFVDVNKDEK----QFMHL 322
              |  :| : |      || | |  :   :  : |:  ::   |: ||:      | |
Db        205 -PPKKIEFLQSCVLEFVDRTSSDLICGAEPYVEGQYRKYNNNSGFVSNDERNTPQSFSHF 263

Qy        323 WNSFVRKQRVIAD----GHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
                  | :| |    |           : : ::| |  :       :  |   |
Db        264 TYEHSNHQLLIIDIQGVG---------DHYTDPQIHTYDGV----GFGIGNLGQKGFEKF 310

Qy        379 ATINNCNTILE--NCRNSSDTTTTNNNNSVDRPS---------DSNTNNNNIVDHPNDIN 427
              : || | :  | :: :  :  ::  :| ||          |:| ||||  :: |: |
Db        311 LDTHKCNAICQYLNLQSINPKSEKSDCGTVPRPDLIFPDTSERDNNNNNNN--NNNNNNN 368

Qy        428 NKNNVDNKDNNSRDKVIK 445
              | ||  | :||:   : |
Db        369 NNNNNSNNNNNNNSSISK 386


RESULT 5
US-08-700-152A-4
; Sequence 4, Application US/08700152A
; Patent No. 5994622
;  GENERAL INFORMATION:
;    APPLICANT:  Jofuku, K. Diane
;    APPLICANT:  Okamuro, Jack K.
;    TITLE OF INVENTION:  Methods for Improving Seeds
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,152A
;      FILING DATE:  20-AUG-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  02307O-067200US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  432 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-700-152A-4

  Query Match             4.3%;  Score 103;  DB 2;  Length 432;
  Best Local Similarity   21.5%;  Pred. No. 0.023;
  Matches   88;  Conservative   54;  Mismatches  134;  Indels  134;  Gaps   20;

Qy         74 DCNNTLQKTEVREDCSCPFC----SMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNV 129
              | |:   :|: ||: |  ||    |   |||       |||
Db          3 DLNDAPHQTQ-REEESEEFCYSSPSKRVGSFS------NSS------------------- 36

Qy        130 SVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTE 189
                :| :  |:||||||:               |   ||       |  :||
Db         37 ----SSAVVIEDGSDDDEL-------------NRVRPNNPLVTHQFFPEMDS-------- 71

Qy        190 NGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTR 249
                  | |    |:| |     :|  :     : |  : :  | |:   |  |  |:|
Db         72 ------NGGGVASGFPRAHWFGVKFCQSDLATGSSAGKATNVAAAVV---EPAQPLKKSR 122

Qy        250 KLSAERSEA-------------RSHLLLQKRQFY-----HSHRVQPMALEQVMSDRDSED 291
              :   ||                ||:    :| |       :|          :  | |
Db        123 RGPRSRSSQYRGVTFYRRTGRWESHIWDCGKQVYLGGFDTAHAAARAYDRAAIKFRGVEA 182

Qy        292 EVDDDVADFED--RQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFY 349
              ::: :: |::|  :||       |  :::|:|:    :|:|              | |
Db        183 DINFNIDDYDDDLKQM------TNLTKEEFVHV----LRRQST-----------GFPRGS 221

Qy        350 EKELHRYSSLFWCWR------LFLIKLWNH-GLVDSATINNCNTILENCR---------N 393
              |  :|  :|  | |       || | : : || |         | :|  |         |
Db        222 SK--YRGVTLHKCGRWEARMGQFLGKKYVYLGLFD--------TEVEAARAYDKAAIKCN 271

Qy        394 SSDTTTTNNNNSVDRP--SDSNTNNNNIVDHPNDINNKNNVDNKDNNSRD 441
                |  | : :  |    ::|: |     ||  |::  |: ::| : |:|
Db        272 GKDAVTNFDPSIYDEELNAESSGNPTTPQDHNLDLSLGNSANSK-HKSQD 320


RESULT 6
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
;  APPLICANT:  Nislow, Corey
;  APPLICANT:  Sakowicz, Roman
;  APPLICANT:  Beraud, Christophe
;  TITLE OF INVENTION:  Antifungal Assay
```

```
;  FILE REFERENCE:  1015
;  CURRENT APPLICATION NUMBER: US/09/541,782
;  CURRENT FILING DATE:  2000-04-03
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 1073
;   TYPE: PRT
;   ORGANISM: S.pombe
US-09-541-782-6

  Query Match             4.3%;  Score 103;  DB 4;  Length 1073;
  Best Local Similarity   18.3%;  Pred. No. 0.1;
  Matches  108;  Conservative   75;  Mismatches  200;  Indels  206;  Gaps   24;

Qy         13 EEVISTDE-------------NLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKR 58
              || ||| |               | |:: | :    :  : | |       ||
Db        394 EETISTLEYAARAKSIRNKPQNNQLVFRKVLIKDLVLDIERLKND------LN------A 441

Qy         59 KRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEF 118
               ||      :    ||:  : :|  ::        |       : |  :: || :  ::
Db        442 TRKKNGVYLAESTYKELMDRVQNKDL-------LCQEQARKLEVLDLNVKSSREQLQYVS 494

Qy        119 KLFEEY---------QTVNVSVKLNSFIFEEEGSDDD---------KFEP---------- 150
              | :|:         | || | :| |   | |    ::         |:|
Db        495 KSNQEHKKEVEALQLQLVNSSTELESVKSENEKLKNELVLEIEKRKKYETNEAKITTVAT 554

Qy        151 -------------FSLCSK--PRKRRQRGGRNNTRRLKVCFLPL---DSPSLTNGTENGI 192
                          || |   :|  :   ||   ||   | :       | |: |
Db        555 DLSQYYRESKEYIASLYEKLDRTERNNKENENNFWNLKFNLLTMLRSFHGSFTDETNGYF 614

Qy        193 TLLNDGNRGLG----------YPEATELAGQFE------MTSNIPPAIAHSSLDAGAKVI 236
              |||||  |  :                |::   |:        ::    |: :||||     :
Db        615 TLLNDFNASMEELLNTHSNQLLISMTKITEHFQSLDEALQSARSSCAVPNSSLD-----L 669

Qy        237 LTSE------AVVPA--------------------TKTRKLSAERSEARSHLLLQKRQFY 270
              : ||      ::: |                    ::  :|  :  |:   |: : |  |
Db        670 IVSELKDSKNSLLDALEHSLQDISMSSQKLGNGISSELIELQKDMKESYRQLVQELRSLY 729

Qy        271 ---HSH-RVQPMALEQVMSDRDS----------------EDEVDDDVADFEDRQMLDDFV 310
                 |:|   |   :  | :| |:                | : |  : || :|  |
Db        730 NLQHTHEESQKELMYGVRNDIDALVKTCTTSLNDADIILSDYISDQKSKFESKQQ-DLIA 788

Qy        311 DVNKDEKQFM----------------HL--WNSFVRKQRVIADG---------------- 336
              :: |    |:                ||   || :||   | :
Db        789 NIGKIVSNFLQEQNESLYTKADILHSHLNDTNSNIRKANEIMNNRSEEFLRNAASQAEIV 848

Qy        337 -----HISWACEAFSRFYE---KELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTIL 388
                  |    | |:  :   | :|  |    :   | | :    |:  :  :   : :|
Db        849 GANKERIQKTVENGSQLLDSKSKAIHSNSRSMYDHCLALAESQKQGV--NLEVQTLDRLL 906

Qy        389 ENCRNSSDTTTT--------------NNNNSVDRPSDSNTNNNNIVDH 422
              :  :  |:  |                ||:| :|     :|    | ||
Db        907 QKVKEHSEDNTKEKHQQLLDLLESLVGNNDNLIDSIKTPHTELQKITDH 955


RESULT 7
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
;  APPLICANT: Nislow, Corey
;  APPLICANT:  Sakowicz, Roman
;  APPLICANT:  Beraud, Christophe
;  TITLE OF INVENTION: Antifungal Assay
;  FILE REFERENCE: 1015
;  CURRENT APPLICATION NUMBER: US/09/723,820
;  CURRENT FILING DATE:  2000-11-28
;  PRIOR APPLICATION NUMBER: 09/541,782
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 1073
;   TYPE: PRT
;   ORGANISM: S.pombe
US-09-723-820-6

  Query Match             4.3%;  Score 103;  DB 4;  Length 1073;
  Best Local Similarity   18.3%;  Pred. No. 0.1;
  Matches  108;  Conservative   75;  Mismatches  200;  Indels  206;  Gaps   24;

Qy         13 EEVISTDE-------------NLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKR 58
              || ||| |               | |:: | :    :  : | |       ||
Db        394 EETISTLEYAARAKSIRNKPQNNQLVFRKVLIKDLVLDIERLKND------LN------A 441

Qy         59 KRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEF 118
               ||      :    ||:  : :|  ::        |       : |  :: || :  ::
Db        442 TRKKNGVYLAESTYKELMDRVQNKDL-------LCQEQARKLEVLDLNVKSSREQLQYVS 494

Qy        119 KLFEEY---------QTVNVSVKLNSFIFEEEGSDDD---------KFEP---------- 150
              | :|:         | || | :| |   | |    ::         |:|
Db        495 KSNQEHKKEVEALQLQLVNSSTELESVKSENEKLKNELVLEIEKRKKYETNEAKITTVAT 554

Qy        151 -------------FSLCSK--PRKRRQRGGRNNTRRLKVCFLPL---DSPSLTNGTENGI 192
                          || |   :|  :   ||   ||   | :       | |: |
Db        555 DLSQYYRESKEYIASLYEKLDRTERNNKENENNFWNLKFNLLTMLRSFHGSFTDETNGYF 614

Qy        193 TLLNDGNRGLG----------YPEATELAGQFE------MTSNIPPAIAHSSLDAGAKVI 236
              |||||  |  :                |::   |:        ::    |: :||||     :
Db        615 TLLNDFNASMEELLNTHSNQLLISMTKITEHFQSLDEALQSARSSCAVPNSSLD-----L 669

Qy        237 LTSE------AVVPA--------------------TKTRKLSAERSEARSHLLLQKRQFY 270
              : ||      ::: |                    ::  :|  :  |:   |: : |  |
Db        670 IVSELKDSKNSLLDALEHSLQDISMSSQKLGNGISSELIELQKDMKESYRQLVQELRSLY 729

Qy        271 ---HSH-RVQPMALEQVMSDRDS----------------EDEVDDDVADFEDRQMLDDFV 310
                 |:|   |   :  | :| |:                | : |  : || :|  |
Db        730 NLQHTHEESQKELMYGVRNDIDALVKTCTTSLNDADIILSDYISDQKSKFESKQQ-DLIA 788

Qy        311 DVNKDEKQFM----------------HL--WNSFVRKQRVIADG---------------- 336
              :: |    |:                ||   || :||   | :
Db        789 NIGKIVSNFLQEQNESLYTKADILHSHLNDTNSNIRKANEIMNNRSEEFLRNAASQAEIV 848

Qy        337 -----HISWACEAFSRFYE---KELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTIL 388
                  |    | |:  :   | :|  |    :   | | :    |:  :  :   : :|
Db        849 GANKERIQKTVENGSQLLDSKSKAIHSNSRSMYDHCLALAESQKQGV--NLEVQTLDRLL 906

Qy        389 ENCRNSSDTTTT--------------NNNNSVDRPSDSNTNNNNIVDH 422
              :  :  |:  |                ||:| :|     :|    | ||
Db        907 QKVKEHSEDNTKEKHQQLLDLLESLVGNNDNLIDSIKTPHTELQKITDH 955


RESULT 8
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
;  APPLICANT:  Beraud, Christophe
;  APPLICANT:  Sakowicz, Roman
;  APPLICANT:  Wood, Kenneth
;  TITLE OF INVENTION:  No. 6355466el motor proteins and methods for
;  TITLE OF INVENTION:  their use
;  FILE REFERENCE:  1017
;  CURRENT APPLICATION NUMBER: US/09/572,191
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1388
;   TYPE: PRT
;   ORGANISM: Human
US-09-572-191-2

  Query Match             4.1%;  Score 98.5;  DB 4;  Length 1388;
  Best Local Similarity   19.0%;  Pred. No. 0.46;
```

```
  Matches   79;  Conservative   81;  Mismatches  150;  Indels  105;  Gaps   18;

Qy         85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF-------- 136
              :| |       |      :|  ||:|   :| |||     : : :  :| |
Db        573 KEPCLFANTEKLKAQLLQIQTELNNSKQEYE-EFKELTRKRQLELESELQSLQKANLNLE 631

Qy        137 -IFE-------EEGSDDDKF--EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
               : |       :| |  :|    |   : : | |  |    |          :| ||  :
Db        632 NLLEATKACKRQEVSQLNKIHAETLKIITTPTKAYQLHSRP---------VPKLSPEM-- 680

Qy        187 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 246
                        |: |  | : : :     :   :|| :   : :| :: : | :  : |
Db        681 -----------GSFGSLYTQNSSILDNDILNEPVPPEMNEQAFEAISEELRTVQEQMSAL 729

Qy        247 KTRKLSAERSEARSHLLLQK---RQFYHSHRVQPMALEQVMSDRDSEDEVDDDV----AD 299
              :  ||     |   :| ||:   :  :|| ::| :   : :     ::|:   :
Db        730 QA-KLD---EEEHKNLKLQQHVDKLEHHSTQMQELFSSERIDWTKQQEELLSQLNVLEKQ 785

Qy        300 FEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVI---ADGHISWACEAFSRF---YEKEL 353
               :: |  :||:              | |   ||:   ||  :|     :| |    |||
Db        786 LQETQTKNDFL------------KSEVHDLRVVLHSADKELSSVKLEYSSFKTNQEKEF 832

Qy        354 HRYSSLFWCWRLFL--IKLWNHGLVDSAT---------------INNCNTILENCRNSS 395
              :: |      :| |  ::| |  |::|                 |:  :  |:| :  :
Db        833 NKLSERHMHVQLQLDNLRLENEKLLESKACLQDSYDNLQEIMKFEIDQLSRNLQNFKKEN 892

Qy        396 DTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNV---DNKDNNSRD--KVIK 445
              :|  :: || ::           :::   : ||| ::   ::|:|:|::  ||::
Db        893 ETLKSDLNNLME----------LLEAEKERNNKLSLQFEEDKENSSKEILKVLE 936


RESULT 9
US-09-723-262-2
; Sequence 2, Application US/09723262
; Patent No. 6379912
; GENERAL INFORMATION:
;  APPLICANT: Beraud, Christophe
;  APPLICANT:  Sakowicz, Roman
;  APPLICANT:  Wood, Kenneth
;  TITLE OF INVENTION: No. 6379912el motor proteins and methods for
;  TITLE OF INVENTION:  their use
;  FILE REFERENCE: 1017
;  CURRENT APPLICATION NUMBER: US/09/723,262
;  CURRENT FILING DATE:  2000-11-27
;  PRIOR APPLICATION NUMBER: US 09/572,191
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1388
;   TYPE: PRT
;   ORGANISM: Human
US-09-723-262-2

  Query Match             4.1%;  Score 98.5;  DB 4;  Length 1388;
  Best Local Similarity   19.0%;  Pred. No. 0.46;
  Matches   79;  Conservative   81;  Mismatches  150;  Indels  105;  Gaps   18;

Qy         85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF-------- 136
              :| |       |      :|  ||:|   :| |||     : : :  :| |
Db        573 KEPCLFANTEKLKAQLLQIQTELNNSKQEYE-EFKELTRKRQLELESELQSLQKANLNLE 631

Qy        137 -IFE-------EEGSDDDKF--EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
               : |       :| |  :|    |   : : | |  |    |          :| ||  :
Db        632 NLLEATKACKRQEVSQLNKIHAETLKIITTPTKAYQLHSRP---------VPKLSPEM-- 680

Qy        187 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 246
                        |: |  | : : :     :   :|| :   : :| :: : | :  : |
Db        681 -----------GSFGSLYTQNSSILDNDILNEPVPPEMNEQAFEAISEELRTVQEQMSAL 729

Qy        247 KTRKLSAERSEARSHLLLQK---RQFYHSHRVQPMALEQVMSDRDSEDEVDDDV----AD 299
              :  ||     |   :| ||:   :  :|| ::| :   : :     ::|:   :
Db        730 QA-KLD---EEEHKNLKLQQHVDKLEHHSTQMQELFSSERIDWTKQQEELLSQLNVLEKQ 785

Qy        300 FEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVI---ADGHISWACEAFSRF---YEKEL 353
               :: |  :||:              | |   ||:   ||  :|     :| |    |||
Db        786 LQETQTKNDFL------------KSEVHDLRVVLHSADKELSSVKLEYSSFKTNQEKEF 832

Qy        354 HRYSSLFWCWRLFL--IKLWNHGLVDSAT---------------INNCNTILENCRNSS 395
              :: |      :| |  ::| |  |::|                 |:  :  |:| :  :
Db        833 NKLSERHMHVQLQLDNLRLENEKLLESKACLQDSYDNLQEIMKFEIDQLSRNLQNFKKEN 892

Qy        396 DTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNV---DNKDNNSRD--KVIK 445
              :|  :: || ::           :::   : ||| ::   ::|:|:|::  ||::
Db        893 ETLKSDLNNLME----------LLEAEKERNNKLSLQFEEDKENSSKEILKVLE 936


RESULT 10
US-09-723-219-2
; Sequence 2, Application US/09723219
; Patent No. 6391613
; GENERAL INFORMATION:
;  APPLICANT: Beraud, Christophe
;  APPLICANT:  Sakowicz, Roman
;  APPLICANT:  Wood, Kenneth
;  TITLE OF INVENTION: No. 6391613el motor proteins and methods for
;  TITLE OF INVENTION:  their use
;  FILE REFERENCE: 1017
;  CURRENT APPLICATION NUMBER: US/09/723,219
;  CURRENT FILING DATE:  2000-11-27
;  PRIOR APPLICATION NUMBER: US 09/572,191
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1388
;   TYPE: PRT
;   ORGANISM: Human
US-09-723-219-2

  Query Match             4.1%;  Score 98.5;  DB 4;  Length 1388;
  Best Local Similarity   19.0%;  Pred. No. 0.46;
  Matches   79;  Conservative   81;  Mismatches  150;  Indels  105;  Gaps   18;

Qy         85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF-------- 136
              :| |       |      :|  ||:|   :| |||     : : :  :| |
Db        573 KEPCLFANTEKLKAQLLQIQTELNNSKQEYE-EFKELTRKRQLELESELQSLQKANLNLE 631

Qy        137 -IFE-------EEGSDDDKF--EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
               : |       :| |  :|    |   : : | |  |    |          :| ||  :
Db        632 NLLEATKACKRQEVSQLNKIHAETLKIITTPTKAYQLHSRP---------VPKLSPEM-- 680

Qy        187 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 246
                        |: |  | : : :     :   :|| :   : :| :: : | :  : |
Db        681 -----------GSFGSLYTQNSSILDNDILNEPVPPEMNEQAFEAISEELRTVQEQMSAL 729

Qy        247 KTRKLSAERSEARSHLLLQK---RQFYHSHRVQPMALEQVMSDRDSEDEVDDDV----AD 299
              :  ||     |   :| ||:   :  :|| ::| :   : :     ::|:   :
Db        730 QA-KLD---EEEHKNLKLQQHVDKLEHHSTQMQELFSSERIDWTKQQEELLSQLNVLEKQ 785

Qy        300 FEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVI---ADGHISWACEAFSRF---YEKEL 353
               :: |  :||:              | |   ||:   ||  :|     :| |    |||
Db        786 LQETQTKNDFL------------KSEVHDLRVVLHSADKELSSVKLEYSSFKTNQEKEF 832

Qy        354 HRYSSLFWCWRLFL--IKLWNHGLVDSAT---------------INNCNTILENCRNSS 395
              :: |      :| |  ::| |  |::|                 |:  :  |:| :  :
Db        833 NKLSERHMHVQLQLDNLRLENEKLLESKACLQDSYDNLQEIMKFEIDQLSRNLQNFKKEN 892

Qy        396 DTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNV---DNKDNNSRD--KVIK 445
              :|  :: || ::           :::   : ||| ::   ::|:|:|::  ||::
Db        893 ETLKSDLNNLME----------LLEAEKERNNKLSLQFEEDKENSSKEILKVLE 936


RESULT 11
US-08-559-896B-2
; Sequence 2, Application US/08559896B
```

```
; Patent No. 6310046
;  GENERAL INFORMATION:
;    APPLICANT:  Patrick E. Duffy
;    APPLICANT:  Christian F. Ockenhouse
;    TITLE OF INVENTION:  SEQUESTRIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John Moran
;      STREET:  USA MRMC -  MCMR-JA
;      CITY:  FORT DETRICK, FREDERICK
;      STATE:  MARYLAND
;      COUNTRY:  USA
;      ZIP:  21702-5012
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh 7.5
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/559,896B
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Moran, John
;      REGISTRATION NUMBER:  26,313
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 619-2065
;      TELEFAX:  (301) 619-7714
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  652 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  Linear
US-08-559-896B-2

  Query Match             4.1%;  Score 98;  DB 4;  Length 652;
  Best Local Similarity   21.8%;  Pred. No. 0.15;
  Matches   41;  Conservative   34;  Mismatches   71;  Indels   42;  Gaps    7;

Qy        265 QKRQFYHSHRVQPMALEQV-MSDRDSEDEV--------DDDVADFEDRQMLDDFVDVNKD 315
              ::::  |  ::  :  |::  |:|  |::         | |      | :|:   |: |:
Db        289 EEKEKIHREKLHKIEKEKINKMDKDQIDKIYEEELNKMDSDEIQHVRRAILE---DIQKE 345

Qy        316 EKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGL 375
              : | : |                   |   | |::|| |       : | : |
Db        346 KIQNLEL--------------------EEIDRLYKEELDRMDR----EARYEIPMRN--- 378

Qy        376 VDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNK 435
                  :     | |  | :| |:       | :|    |:| :|||  ::  |:|| ||  :
Db        379 --LSRNEKDNIIHRNIKNESNQKNKKENVNVFIIHDNNDSNNNNNNNNRDVNNLNN-KHT 435

Qy        436 DNNSRDKV 443
              :||  : |
Db        436 NNNYNENV 443


RESULT 12
US-08-446-855A-2
; Sequence 2, Application US/08446855A
; Patent No. 5849573
;  GENERAL INFORMATION:
;    APPLICANT:  Stewart, Thomas S
;    APPLICANT:  Flores, Maria V
;    APPLICANT:  O'Sullivan, William J
;    TITLE OF INVENTION:  Nucleotide sequence encoding carbamoyl
;    TITLE OF INVENTION:  phosphate synthetase II
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye PC
;      STREET:  1100 No. 5849573th Glebe Road, 8th Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,855A
;      FILING DATE:  06-Jul-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mitchard, Leonard C
;      REGISTRATION NUMBER:  29.009
;      REFERENCE/DOCKET NUMBER:  47-80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2391 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-446-855A-2

  Query Match             4.0%;  Score 94;  DB 2;  Length 2391;
  Best Local Similarity   21.4%;  Pred. No. 3.4;
  Matches   44;  Conservative   40;  Mismatches   64;  Indels   58;  Gaps   12;

Qy        266 KRQFYHSHRVQPMALEQVMSDRDSED---EVDDDVADFEDRQMLDDFVDVNKDEKQFMHL 322
              | :| :::       | : :|   ||   |::  :::| :   :  | | :: :    |
Db        269 KEEFNYTN-------EMITNDSSMEDHDNEINGSISNFNNCPSISSF-DKSESKNVINH- 319

Qy        323 WNSFVR-KQRVIADGHISWACEAFSRFYEKELH---------RYSSLFWCWRLFLIKLWN 372
                : :| |  :|            |  | |:||         :  | |   :|: |  ::
Db        320 --TLLRDKMNLITS----------SEEYLKDLHNCNFSNSSDKNDSFF---KLYGICEYD 364

Qy        373 HGLVD-----SATINNCN-----TILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNI--- 419
              |:|      |   || :      : :|    |:     |||: :  :: | ||| :
Db        365 KYLIDLEENASFHYNNVDEYGYYDVNKNTNILSNNKIEQNNNNENNKNNKNNNNNEVDYI 424

Qy        420 -VDHPNDINNK-------NNVDNKDN 437
               |  |::|:|       ||  | ::
Db        425 KKDEDNNVNSKVFYSQYNNNAQNNEH 450


RESULT 13
US-09-150-741-2
; Sequence 2, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
;  APPLICANT: Stewart et al.
;  TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
;  TITLE OF INVENTION: Synthetase II
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/150,741
;  CURRENT FILING DATE: 1998-09-10
;  EARLIER APPLICATION NUMBER: PL6380
;  EARLIER FILING DATE: 1992-12-16
;  EARLIER APPLICATION NUMBER: AU93/00617
;  EARLIER FILING DATE: 1993-12-02
;  EARLIER APPLICATION NUMBER: 08/446,855
;  EARLIER FILING DATE: 1995-07-06
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2391
```

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2

  Query Match             4.0%;  Score 94;  DB 4;  Length 2391;
  Best Local Similarity   21.4%;  Pred. No. 3.4;
  Matches   44;  Conservative   40;  Mismatches   64;  Indels   58;  Gaps   12;

Qy        266 KRQFYHSHRVQPMALEQVMSDRDSED---EVDDDVADFEDRQMLDDFVDVNKDEKQFMHL 322
              | :| :::       | : :|   ||   |::  :::| :   :  | | :: :   |
Db        269 KEEFNYTN-------EMITNDSSMEDHDNEINGSISNFNNCPSISSF-DKSESKNVINH- 319

Qy        323 WNSFVR-KQRVIADGHISWACEAFSRFYEKELH---------RYSSLFWCWRLFLIKLWN 372
                : :| |  :|           |  | |:||         :  | |   :|: |  ::
Db        320 --TLLRDKMNLITS----------SEEYLKDLHNCNFSNSSDKNDSFF---KLYGICEYD 364

Qy        373 HGLVD-----SATINNCN----TILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNI--- 419
              |:|     |   || :      : :|    |:    |||: :  :: | ||| :
Db        365 KYLIDLEENASFHYNNVDEYGYYDVNKNTNILSNNKIEQNNNNENNKNNKNNNNNEVDYI 424

Qy        420 -VDHPNDINNK-------NNVDNKDN 437
               |  |::|:|       ||  | ::
Db        425 KKDEDNNVNSKVFYSQYNNNAQNNEH 450


RESULT 14
US-09-384-162-6
; Sequence 6, Application US/09384162
; Patent No. 6376747
; GENERAL INFORMATION:
;  APPLICANT: Xing, Ti
;  APPLICANT: Malik, Kamal
;  APPLICANT: Martin-Heller, Teresa
;  APPLICANT: Miki L., Brian
;  TITLE OF INVENTION: No. 6376747el Plant-Derived Map Kinase Kinase
;  FILE REFERENCE: 08-884280US
;  CURRENT APPLICATION NUMBER: US/09/384,162
;  CURRENT FILING DATE: 1999-08-27
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 286
;   TYPE: PRT
;   ORGANISM: Dictyostelium discoideum
US-09-384-162-6

  Query Match             3.9%;  Score 92;  DB 4;  Length 286;
  Best Local Similarity   37.0%;  Pred. No. 0.18;
  Matches   20;  Conservative   10;  Mismatches   14;  Indels   10;  Gaps    2;

Qy        381 INNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDN 434
              ||| |  : |  |:::    ||||      ::| ||||::    ||:|   ||:
Db        195 INNSNNNIRNSNNNNNNNNNNNNN------NNNNNNNNVL----DISNGGLVDS 238


RESULT 15
US-09-177-249-2
; Sequence 2, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
;  APPLICANT: Fischer, Robert L.
;  APPLICANT: Ohad, Nir
;  APPLICANT: Kiyosue, Tomohiro
;  APPLICANT: Yadegari, Ramin
;  APPLICANT: Margossian, Linda
;  APPLICANT: Harada, John
;  APPLICANT: Goldberg, Robert B.
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
;  TITLE OF INVENTION: Development in Plants
;  FILE REFERENCE: 023070-086120US
;  CURRENT APPLICATION NUMBER: US/09/177,249
;  CURRENT FILING DATE: 1998-10-22
;  EARLIER APPLICATION NUMBER: US 09/071,838
;  EARLIER FILING DATE: 1998-05-01
;  NUMBER OF SEQ ID NOS: 324
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 689
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-177-249-2

  Query Match             3.8%;  Score 91.5;  DB 4;  Length 689;
  Best Local Similarity   20.2%;  Pred. No. 0.84;
  Matches   49;  Conservative   39;  Mismatches   61;  Indels   93;  Gaps   11;

Qy        246 TKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQM 305
              ||: :| ||     |  :: ||| |:   :   ||| : |: | ||| :|:
Db        142 TKSSQLMAE-----SDSVIGKRQIYY---LNGEALE-LSSEEDEEDEEEDE--------- 183

Qy        306 LDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWAC-----------EAFSRFYEKE- 352
                   :: |:: :|    : |:            |               | ::: | :
Db        184 ----EEIKKEKCEFSEDVDRFI-----------WTVGQDYGLDDLVVRRALAKYLEVDV 227

Qy        353 ---LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTIL---------ENCRN------- 393
                | ||: |          || | |    |:     ||           :||
Db        228 SDILERYNEL---------KLKNDGTAGEASDLTSKTITTAFQDFADRRHCRRCMIFDCH 278

Qy        394 -----SSDTTTTNNNNSVDRPSD------------SNTNNNNIVDHPNDINNKNNVDNK 435
                    :: :: : :|:    |              | |  ::::|: | |:|| | :
Db        279 MHEKYEPESRSSEDKSSLFEDEDRQPCSEHCYLKVRSVTEADHVMDNDNSISNKIVVSDP 338

Qy        436 DN 437
              :|
Db        339 NN 340


Search completed: June 13, 2003, 15:45:55
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:34:30 ; Search time 25 Seconds
(without alignments)
1903.099 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	225	9.5	388	10	US-09-764-864-931	Sequence 931, App
2	221	9.3	739	9	US-09-874-162A-5	Sequence 5, Appli
3	221	9.3	776	9	US-09-874-162A-8	Sequence 8, Appli
4	153	6.4	289	10	US-09-764-864-911	Sequence 911, App
5	127.5	5.4	961	10	US-09-801-368-132	Sequence 132, App
6	126.5	5.3	2150	9	US-10-135-322-17	Sequence 17, Appl
7	125.5	5.3	1501	10	US-09-924-154-17	Sequence 17, Appl
8	124	5.2	1288	9	US-10-006-780-2	Sequence 2, Appli
9	122.5	5.2	1331	9	US-10-087-464-50	Sequence 50, Appl
10	115	4.8	175	10	US-09-764-864-1355	Sequence 1355, Ap
11	113	4.8	807	9	US-09-820-843A-108	Sequence 108, App
12	111.5	4.7	1143	10	US-09-924-154-14	Sequence 14, Appl
13	105	4.4	732	9	US-09-832-292-12	Sequence 12, Appl
14	105	4.4	732	10	US-09-994-485-8	Sequence 8, Appli
15	104.5	4.4	97	9	US-09-858-935B-77	Sequence 77, Appl
16	104.5	4.4	666	10	US-09-801-368-36	Sequence 36, Appl
17	101	4.2	534	10	US-09-801-368-124	Sequence 124, App
18	100.5	4.2	97	9	US-09-858-935B-72	Sequence 72, Appl
19	99.5	4.2	292	10	US-09-764-864-1369	Sequence 1369, Ap
20	99.5	4.2	861	9	US-09-820-843A-109	Sequence 109, App
21	98	4.1	97	9	US-09-858-935B-76	Sequence 76, Appl
22	98	4.1	652	10	US-09-351-794A-2	Sequence 2, Appli
23	97.5	4.1	97	9	US-09-858-935B-73	Sequence 73, Appl
24	97.5	4.1	97	9	US-09-858-935B-74	Sequence 74, Appl
25	97.5	4.1	97	9	US-09-858-935B-75	Sequence 75, Appl
26	97	4.1	440	9	US-10-083-357-1266	Sequence 1266, Ap
27	97	4.1	559	10	US-09-801-368-256	Sequence 256, App
28	96.5	4.1	97	9	US-09-858-935B-70	Sequence 70, Appl
29	96.5	4.1	97	9	US-09-858-935B-71	Sequence 71, Appl
30	95	4.0	1093	10	US-09-801-368-392	Sequence 392, App
31	94.5	4.0	393	9	US-10-153-668-76	Sequence 76, Appl
32	94.5	4.0	393	9	US-10-153-668-78	Sequence 78, Appl
33	94.5	4.0	556	9	US-10-012-896-1005	Sequence 1005, Ap
34	94.5	4.0	1610	9	US-10-155-533-9	Sequence 9, Appli
35	93.5	3.9	497	9	US-09-820-843A-32	Sequence 32, Appl
36	93	3.9	676	10	US-09-801-368-302	Sequence 302, App
37	93	3.9	2086	10	US-09-815-242-5639	Sequence 5639, Ap
38	93	3.9	5795	10	US-09-815-242-12610	Sequence 12610, A
39	92.5	3.9	704	10	US-09-801-368-218	Sequence 218, App
40	92.5	3.9	758	10	US-09-801-368-224	Sequence 224, App
41	92.5	3.9	762	9	US-09-738-626-4825	Sequence 4825, Ap
42	91.5	3.8	91	9	US-09-858-935B-69	Sequence 69, Appl
43	91.5	3.8	689	10	US-09-071-838-2	Sequence 2, Appli
44	91.5	3.8	1786	9	US-09-742-096-3	Sequence 3, Appli
45	91	3.8	1051	10	US-09-911-888-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-764-864-931
; Sequence 931, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 931
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-931

Query Match 9.5%; Score 225; DB 10; Length 388;
Best Local Similarity 21.1%; Pred. No. 6.2e-13;
Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

```
Qy          58 RKRK-----SRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
                 || |     :|    : : :   ||| |:|| |:|  ||:|:: |     |  ||   |
Db          62 RKEKDTPNENRQKLRIFYQFLYNNNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHS 121

Qy         113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN 168
               | | :    :  ::||:         || |       |    |:  || |    ||
Db         122 RFIFNYVYHPKGARIDVSI--------NEGYDG------SYAGNPQDIHRQPGFAFSRNG 167

Qy         169 -TRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHS 227
                :|  :  : : :  | |  : :      ||
Db         168 PVKRTPITHILVCRPKRTKASMSEFLESEDG---------------------------- 198

Qy         228 SLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDR 287
                                  ::  :|: :  |     | ::|     |:  :::  :
Db         199 ----------------------EVEQQRTYSSGH----NRLYFHRDTCLPLRPQEM--EV 230

Qy         288 DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSR 347
               |||||  | :    :    :::| |||: ||: | |||  | |   |||  :: ||  |
Db         231 DSEDEKDPEWLREKTITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVE 290
```

```
Qy        348 FYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              | ::: :  :|   : | |: : :  |:   :|:   | |   :           ::
Db  .     291 NYGQKIIK-KNLCRNFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLE 338

Qy        408 RPSDSNTNNNNIVDHPN-------DINNKNNVDNKDNNS 439
              :   ::  |  | :  |       :||:|      |: |
Db        339 KGESASPANEEITEEQNGTANGFSEINSKEKALETDSVS 377


RESULT 2
US-09-874-162A-5
; Sequence 5, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
;  APPLICANT: Koontz, Jason
;  APPLICANT:  Sklar, Jeffrey
;  TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
;  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
;  FILE REFERENCE: 05311-024001
;  CURRENT APPLICATION NUMBER: US/09/874,162A
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,093
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 739
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-874-162A-5

  Query Match             9.3%;  Score 221;  DB 9;  Length 739;
  Best Local Similarity   21.1%;  Pred. No. 3.8e-12;
  Matches   84;  Conservative   64;  Mismatches  151;  Indels  100;  Gaps   12;

Qy         58 RKRK-----SRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
              || |     :|    : : :   ||| |:|| |:|  ||:|:: |     |  ||   |
Db        413 RKEKDTPNENRQKLRIFYQFLYNNNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHS 472

Qy        113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN 168
               | | :    :   ::||:         |  |       |    |:   || |    ||
Db        473 RFIFNYVYHPKGARIDVSI--------NECYDG------SYAGNPQDIHRQPGFAFSRNG 518

Qy        169 -TRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHS 227
               :| : : :  |  | : :      ||
Db        519 PVKRTPITHILVCRPKRTKASMSEFLESEDG---------------------------- 549

Qy        228 SLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDR 287
                                  ::  :|: :  |     | ::||    |:  ::: :
Db        550 ----------------------EVEQQRTYSSGH----NRLYFHSDTCLPLRPQEM--EV 581

Qy        288 DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSR 347
              ||||| | :    :    :::| ||| ||: | |||  | |   |||  :: ||  |
Db        582 DSEDEKDPEWLREKTITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVE 641

Qy        348 FYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              | ::: :  :|   : | |: : :  |:   :|:   | |   :           ::
Db        642 NYGQKIIK-KNLCRNFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLE 689

Qy        408 RPSDSNTNNNNIVDHPN-------DINNKNNVDNKDNNS 439
              :   ::  |  | :  |       :||:|      |: |
Db        690 KGESASPANEEITEEQNGTANGFSEINSKEKALETDSVS 728


RESULT 3
US-09-874-162A-8
; Sequence 8, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
;  APPLICANT: Koontz, Jason
;  APPLICANT:  Sklar, Jeffrey
;  TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
;  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
;  FILE REFERENCE: 05311-024001
;  CURRENT APPLICATION NUMBER: US/09/874,162A
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,093
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 776
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-874-162A-8

  Query Match             9.3%;  Score 221;  DB 9;  Length 776;
  Best Local Similarity   21.1%;  Pred. No. 4.1e-12;
  Matches   84;  Conservative   64;  Mismatches  151;  Indels  100;  Gaps   12;

Qy         58 RKRK-----SRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
              || |     :|    : : :   ||| |:|| |:|  ||:|:: |     |  ||   |
Db        450 RKEKDTPNENRQKLRIFYQFLYNNNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHS 509

Qy        113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN 168
               | | :    :   ::||:         |  |       |    |:   || |    ||
Db        510 RFIFNYVYHPKGARIDVSI--------NECYDG------SYAGNPQDIHRQPGFAFSRNG 555

Qy        169 -TRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHS 227
               :| : : :  |  | : :      ||
Db        556 PVKRTPITHILVCRPKRTKASMSEFLESEDG---------------------------- 586

Qy        228 SLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDR 287
                                  ::  :|: :  |     | ::||    |:  ::: :
Db        587 ----------------------EVEQQRTYSSGH----NRLYFHSDTCLPLRPQEM--EV 618

Qy        288 DSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSR 347
              ||||| | :    :    :::| ||| ||: | |||  | |   |||  :: ||  |
Db        619 DSEDEKDPEWLREKTITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVE 678

Qy        348 FYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              | ::: :  :|   : | |: : :  |:   :|:   | |   :           ::
Db        679 NYGQKIIK-KNLCRNFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLE 726

Qy        408 RPSDSNTNNNNIVDHPN-------DINNKNNVDNKDNNS 439
              :   ::  |  | :  |       :||:|      |: |
Db        727 KGESASPANEEITEEQNGTANGFSEINSKEKALETDSVS 765


RESULT 4
US-09-764-864-911
; Sequence 911, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 911
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-864-911

  Query Match             6.4%;  Score 153;  DB 10;  Length 289;
  Best Local Similarity   23.9%;  Pred. No. 2.7e-06;
  Matches   49;  Conservative   43;  Mismatches   89;  Indels   24;  Gaps    5;

Qy        245 ATKTRKLSAERSEA---RSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFE 301
              |: :  | :|  |    |::     | ::||    |:  :::  : |||||  | :   :
Db         88 ASMSEFLESEDGEVEQQRTYSSGHNRLYFHSDTCLPLRPQEM--EVDSEDEKDPEWLREK 145
```

```
Qy        302 DRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFW 361
              :::| |||: ||: | |||  | |   |||  :: ||  |   | ::: :  :|
Db        146 TITQIEEFSDVNEGEKEVMKLWNLHVMKHGFIADNQMNHACMLFVENYGQKIIK-KNLCR 204

Qy        362 CWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVD 421
               : | |: : :  |:   :|:   | |   :             :::   ::  |  | :
Db        205 NFMLHLVSMHDFNLISIMSIDKAVTKLREMQ----------QKLEKGESASPANEEITE 253

Qy        422 HPN-------DINNKNNVDNKDNNS 439
                |       :||:|      |: |
Db        254 EQNGTANGFSEINSKEKALETDSVS 278


RESULT 5
US-09-801-368-132
; Sequence 132, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132
;   LENGTH: 961
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-132

  Query Match             5.4%;  Score 127.5;  DB 10;  Length 961;
  Best Local Similarity   41.3%;  Pred. No. 0.0041;
  Matches   26;  Conservative   13;  Mismatches   23;  Indels    1;  Gaps    1;

Qy        382 NNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINN-KNNVDNKDNNSR 440
              || |   :|  |:|:    ||||: :  :::| |||| :: |: || ||||||  : |
Db        501 NNNNNDNDNDNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSNNIKNNVDNNNTNPA 560

Qy        441 DKV 443
              | :
Db        561 DNI 563


RESULT 6
US-10-135-322-17
; Sequence 17, Application US/10135322
; Patent No. US20020173017A1
; GENERAL INFORMATION:
;  APPLICANT: BENFEY, PN
;  APPLICANT:  HELARIUTTA, Y
;  APPLICANT:  MAHONEN, AP
;  APPLICANT:  BONKE, AWM
;  APPLICANT:  KAUPPINEN, L
;  APPLICANT:  RIIKONEN, M
;  TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
;  FILE REFERENCE: 5914-086-999
;  CURRENT APPLICATION NUMBER: US/10/135,322
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: 60/253,739
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 2150
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-135-322-17

  Query Match             5.3%;  Score 126.5;  DB 9;  Length 2150;
  Best Local Similarity   17.5%;  Pred. No. 0.017;
  Matches   85;  Conservative   64;  Mismatches  175;  Indels  163;  Gaps   16;

Qy          2 CRQNCRAKS----SPEEVISTDENLLIYCKPVRLYNIFHLRSLGN-----PSFLPRCLNY 52
              |  || :      ||  : | :  : :  : :      || :| |     |   || :|:
Db        353 CGSNCGSNGGIPLSPRNLSSLNSGVNVSPRNI------HLNNLNNNSSNLPPLSPRHINF 406

Qy         53 KIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
               |       : :  :   |  : ||:                                |:|::
Db        407 HINVSNLNNNNNNNINPNN--NPNNS----------------------------NNSNN 435

Qy        113 LFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGR------ 166
                          ||| : |:      ||:     |         |   ||
Db        436 --------------NVSPRNNNHNISPRGSN---ISP---------RSNNGGSTTISPR 468

Qy        167 ---------NNTRRLKVCFLPLDSPSL--TNGTENGITLLNDGNRGLGYPEA-TELAGQF 214
                      ||     :   | :|| |   | | :   |    |  |    : |
Db        469 NISNNNNIINNINNNNILTPPRNSPRLENVNPTNSPRLLATSLNSTLPIVSSLTSSNNNN 528

Qy        215 EMTSNIPPAIAHSS------LDAGAKVILTSEAVV--PATKTRKLSAERSEARSHLLLQK 266
              :  :|  |:| :::      :   :: || :: ||          :   |   :: :
Db        529 QSNNNTNPSINNNNGRNGHCIQTISEEILGNKPVVYNNGNNNNNNNTNNSTTSNNNITTN 588

Qy        267 RQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQML--DDFVD----------VNK 314
                ::: :    |                 : : |||   |   || :           ||
Db        589 NNNNNNNNININNVLSTPRKRTKGNHSKTNSLQDFETSSMNGGDDSISGAGSGGSLRRRNK 648

Qy        315 DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHG 374
              |:                                                      | |
Db        649 DDND----------------------------------------------------ENDG 656

Qy        375 LVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDN 434
              ::   || |    |  |||:    |:||: :  :::| ||||  :: |: || || :|
Db        657 NSNNTNSNNSNNNNNNNNNSSNNNNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 716

Qy        435 KDNNSRD 441
              :||: :
Db        717 NNNNNNN 723


RESULT 7
US-09-924-154-17
; Sequence 17, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
;  APPLICANT: Narum, David L.
;  APPLICANT:  Sim, Kim L.
;  TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
;  FILE REFERENCE: 05213-0465  43170-262105
;  CURRENT APPLICATION NUMBER: US/09/924,154
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: US 60/223,525
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 1501
;   TYPE: PRT
;   ORGANISM: Mammalian
```

```
US-09-924-154-17

  Query Match             5.3%;  Score 125.5;  DB 10;  Length 1501;
  Best Local Similarity   19.2%;  Pred. No. 0.012;
  Matches   83;  Conservative   59;  Mismatches  164;  Indels  127;  Gaps   13;

Qy          48 RCLNYKIGAKRKRKSRS----------TGMVVFNYKDCNNTLQKTEV--------REDCS 89
               | | |:| : :  | |          |  :  | |:|:| :  |::         :| ||
Db         670 RKLAYEIQSVKYDKDRKLFSLAKDKNVTTFLKENAKNCSN-IDFTKIFDQLDKLFKERCS 728

Qy          90 CPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVN-VSVKLNSFIFEEEGSDDDKF 148
               |    :|    | :   ::| |  :   |  ||   || : |   | | | |:|
Db         729 CMDTQVLEVKNKEMLSIDSNSEDATDISEKNGEEELYVNHNSVSVASGNKEIEKSKDEK- 787

Qy         149 EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEAT 208
                    :| |  ::                    ||||   :    | :|  |   ||
Db         788 -------QPEKEAKQ--------------------TNGTLT-VRTDKDSDRNKGKDTAT 818

Qy         209 ELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQ 268
               :     | | |  :       : |     :| :     | |: || :    |: :
Db         819 D-------TKNSPENLKVQEHGTNGETIKEEPPKLPESSETLQSQEQLEAEAQKQKQEEE 871

Qy         269 FYHSHRVQPMALEQVMSDRDSE--DEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSF 326
                    :|    ::    |: |   | :::    |: | : |      |:   :
Db         872 PKKKQEEEPKKKQEEEQKREQEQKQEQEEEEQKQEEEQQIQDQSQSGLDQSSKV----GV 927

Qy         327 VRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNT 386
               :|   |: |
Db         928 ASEQNEISSGQ---------------------------------------------------- 938

Qy         387 ILENCRNSS------DTTTTNNNNSVDRPSDSNTNNNNIVDH--------PNDINNKNNV 432
                 :| ::||      :||: | ::   : | :  | |::||          |  ::|:|
Db         939 -EQNVKSSSPEVVPQETTSENGSSQDTKISSTEPNENSVVDRATDSMNLDPEKVHNENMS 997

Qy         433 DNKDNNSRDKVIK 445
               |   |   | :|
Db         998 DPNTNTEPDASLK 1010


RESULT 8
US-10-006-780-2
; Sequence 2, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
;  APPLICANT: Sakowicz, Roman
;  APPLICANT:  Beraud, Christophe
;  APPLICANT:  Guo, Jun
;  APPLICANT:  Freedman, Richard
;  TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
;  TITLE OF INVENTION:  METHODS FOR ITS USE
;  FILE REFERENCE: CYTOP083
;  CURRENT APPLICATION NUMBER: US/10/006,780
;  CURRENT FILING DATE:  2001-11-30
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1288
;   TYPE: PRT
;   ORGANISM: Plasmodium falciparum
US-10-006-780-2

  Query Match             5.2%;  Score 124;  DB 9;  Length 1288;
  Best Local Similarity   21.4%;  Pred. No. 0.014;
  Matches   78;  Conservative   57;  Mismatches  129;  Indels  100;  Gaps   17;

Qy         110 SHDLFEFEFKLFEEYQTVNVSVKLNSFIF-----EEEGSDDDKFEPFSLCSKPRKRRQRG 164
               || :  : |       :| :  |     |  | |:|              |: |
Db         215 SHAILNIDLK------DINKNTSLGKIAFIDLAGSERGAD---------TVSQNKQTQTD 259

Qy         165 GRNNTR---RLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATEL-----AGQFE- 215
               | |  |    || |   :||       :| |   :          | |::      |: :
Db         260 GANINRSLLALKECIRAMDS------DKNHIPFRDS--------ELTKVLRDIFVGKSKS 305

Qy         216 -MTSNIPPAIA----------HSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLL 264
                | :|| | |:          :||      |   |        | | ::|   |:
Db         306 IMIANISPTISCCEQTLNTLRYSSRVKNFKNKSTCINEEDDTNTERISILDSK------- 358

Qy         265 QKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWN 324
                        :   ::| |:                    :   :| :    | : |      |
Db         359 -------GSEMNASSIENVV--------------IKSNHLLSN----NNNNKINRGKIN 392

Qy         325 SFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVD-SATINN 383
                 : :  :: :        | |:  : |    ||||    ::   |    ||::  :|: |
Db         393 DKIERNNILKNKSFDKPREGFTSTFGK----YSSLNDIDKIKKNK--KKGLINYKSTLYN 446

Qy         384 CNTILE----NCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDH--PNDINNKNNVDNKDN 437
                |||  :    |  |::|    ||:|: :  :||::  ||:::|   |:|||  ||:| :|
Db         447 DNTINKKHNNNNNNNNDNNDNNNDNNNNNNNDSSSMVNNMINHMINNNINNNINVNNNNN 506

Qy         438 NSRD 441
               |: :
Db         507 NNNN 510


RESULT 9
US-10-087-464-50
; Sequence 50, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
;  APPLICANT: Chishti, Athar
;  APPLICANT:  Oh, Steven
;  APPLICANT:  Liu, David
;  APPLICANT:  Goel, Vikas
;  APPLICANT:  Li, Xuerong
;  TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses There
;  FILE REFERENCE: S1237/7019
;  CURRENT APPLICATION NUMBER: US/10/087,464
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 06/272,930
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 59
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 1331
;   TYPE: PRT
;   ORGANISM: Plasmodium falciparum
US-10-087-464-50

  Query Match             5.2%;  Score 122.5;  DB 9;  Length 1331;
  Best Local Similarity   39.2%;  Pred. No. 0.02;
  Matches   29;  Conservative   12;  Mismatches   30;  Indels    3;  Gaps    2;

Qy         371 WNHGLVDSATINN--CNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPN-DIN 427
               :|:|  ||   ||    :   |  |: |   ||||  :  :::| ||||  :: |   |
Db         595 YNNGYNDSTDNNNGYNSNSSYNSNNNEDDNNNNNNNDENCDNNNNHNNNNYNNNNNYGNN 654

Qy         428 NKNNVDNKDNNSRD 441
               | || :|||||: |
Db         655 NNNNNNNKDNNNND 668


RESULT 10
US-09-764-864-1355
; Sequence 1355, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1355
;   LENGTH: 175
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (4)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (136)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (166)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1355

  Query Match             4.8%;  Score 115;  DB 10;  Length 175;
  Best Local Similarity   20.4%;  Pred. No. 0.0053;
  Matches   46;  Conservative   30;  Mismatches   74;  Indels   76;  Gaps    8;

Qy         76 NNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNS 135
              ||| |:|| |:|  ||:|:: |     |  ||   |  | | :   :   ::||:
Db          8 NNTRQQTEARDDLHCPWCTLNCRKLYSLLKHLKLCHSRFIFNYVYHPKGARIDVSI---- 63

Qy        136 FIFEEEGSDDDKFEPFSLCSKPRK-RRQRG---GRNN-TRRLKVCFLPLDSPSLTNGTEN 190
                  |  |       |    |:  || |    ||   :| :  : :  | |  : :
Db         64 ----NECYDG------SYAGNPQDIHRQPGFAFSRNGPVKRTPITHILVCRPKRTKASMS 113

Qy        191 GITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRK 250
                 ||                                                     :
Db        114 EFLESEDG----------------------------------------------------E 122

Qy        251 LSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDD 296
              :  :|: :  |     | ::||    |:  :::  : |||| | :
Db        123 VEQQRTYSSGH----NRXYFHSDTCLPLRPQEM--EVDSEDEKDPE 162


RESULT 11
US-09-820-843A-108
; Sequence 108, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
;  APPLICANT: Council of Scientific and Industrial Research
;  TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
;  TITLE OF INVENTION:  USEFUL AS ANTI-INFECTIVES
;  FILE REFERENCE: Q63915
;  CURRENT APPLICATION NUMBER: US/09/820,843A
;  CURRENT FILING DATE:  2001-03-30
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
;   LENGTH: 807
;   TYPE: PRT
;   ORGANISM: Plasmodium falciparum
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: hypothetical protein
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: gi|3845292
US-09-820-843A-108

  Query Match             4.8%;  Score 113;  DB 9;  Length 807;
  Best Local Similarity   35.3%;  Pred. No. 0.076;
  Matches   24;  Conservative   14;  Mismatches   24;  Indels    6;  Gaps    1;

Qy        372 NHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNN 431
              |:||    |:   | | :   |: : || || |::   :: | | |      |: |||||
Db        157 NYGLKKKITLLKRNDIKDEGYNNENITTLNNKNNLKNNNNYNDNRN------NNNNNKNN 210

Qy        432 VDNKDNNS 439
              ::| :||:
Db        211 INNNNNNN 218


RESULT 12
US-09-924-154-14
; Sequence 14, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
;  APPLICANT: Narum, David L.
;  APPLICANT:  Sim, Kim L.
;  TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
;  FILE REFERENCE: 05213-0465  43170-262105
;  CURRENT APPLICATION NUMBER: US/09/924,154
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: US 60/223,525
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 1143
;   TYPE: PRT
;   ORGANISM: Mammalian
US-09-924-154-14

  Query Match             4.7%;  Score 111.5;  DB 10;  Length 1143;
  Best Local Similarity   18.6%;  Pred. No. 0.17;
  Matches   94;  Conservative   70;  Mismatches  171;  Indels  171;  Gaps   24;

Qy          2 CRQNCR--------AKSSPEEVI---STDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCL 50
              |:  |           ||  |: :   : |:|  :|       ||  ::
Db        664 CKNKCSEYKKWIDLKKSEYEKQVDKYTKDKNKKMY------DNIDEVK------------ 705

Qy         51 NYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLN-- 108
              | :     | ||:    | |: | | :    |  : |        |   |    :||
Db        706 NKEANVYLKEKSKECKDVNFDDKIFNES--PNEYEDMCK------KCDEIK----YLNEI 753

Qy        109 ----SSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDD------------KFEPF 151
                  : ||::: :   | :          : :|: | |::   |                ||
Db        754 KYPKTKHDIYDID--TFSDTFGDGTPISINANINEQQSGKDTSNTGNSETSDSPVSHEPE 811

Qy        152 S--------LCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLG 203
              |        |          ||            | :: ||:||
Db        812 SDAAINVEKLSGDESSSETRG----------ILDINDPSVTNNVNE------------- 847

Qy        204 YPEATELAGQFEMTSNIPPAIAHSSLD--AGAKVILTSEAVVPATKTRKLSAERSEARSH 261
                :|:   |    ||:|    : |||:    |: |    :    :  :: |:  |:
Db        848 VHDASNTQGSVSNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQENSSHSSDNSGS 907

Qy        262 LLL-------QKRQFYHS---HRVQPMALEQVMSD------------RDSEDEVDDDVAD 299
              | :         : | |   ||  | ||  : ||            ||||:   |  ::
Db        908 LTIGQVPSEDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSENGRGDTTSN 967

Qy        300 FEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSL 359
              | :       ::  : :   |    |:|      : |                 |:|
Db        968 THDVRR----TNIVSERRVNSH---DFIRNGMANNNAH---------------HQY--- 1001

Qy        360 FWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNI 419
                    : :: |:|::                    : :  |: |  | |  || :: |
Db       1002 -------ITQIENNGII---------------RGQEESAGNSVNYKDNPKRSNFSSEN- 1037

Qy        420 VDHPNDINNKNNVDNKDNNSRDKVIK 445
               ||  :|   |: | |    |:::||
Db       1038 -DHKKNIQEYNSRDTK--RVREEIIK 1060


RESULT 13
US-09-832-292-12
; Sequence 12, Application US/09832292
; Patent No. US20020177205A1
; GENERAL INFORMATION:
;  APPLICANT: Ryazanov, Alexey
;  TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND
;  TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC USES THEREOF
;  FILE REFERENCE: 601-1-098CIP
;  CURRENT APPLICATION NUMBER: US/09/832,292
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: 09/632,131
```

```
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 732
;   TYPE: PRT
;   ORGANISM: Dictyostelium discoideum
US-09-832-292-12

  Query Match             4.4%;  Score 105;  DB 9;  Length 732;
  Best Local Similarity   20.1%;  Pred. No. 0.38;
  Matches   76;  Conservative   61;  Mismatches  167;  Indels   74;  Gaps   17;

Qy        117 EFKLFE-EYQTVNVSVKLNSFIFEEEG---SDDDKFEPFSLCSKPRKRRQRGGRNNTRRL 172
              | |:|   | |    :::   :    |   : | ||| |   :     : : |     |:
Db         34 EGKIFRFPYGTSQTDLQIGKMLPSGSGGGATADSKFEKFKARNTLADIQYKVGDTLYVRV 93

Qy        173 KV------CFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNI----PP 222
              |       ||  : :  :|:|  |     |       |  |    |:  |: :    |
Db         94 KKSKPTNDSLLPTLNIAFLDGSERAIKWEYD-------PYTT--TAQWTCTATLVKVEPV 144

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQF--------YHSHR 274
              |  :      :  |::       |:  : :  |:      | |          |:| :
Db        145 PFAEQAFRKAYHTLDLSKSGASGRYVSKIGKKPTPRPSYFEDVKMQMIAKKWADKYNSFK 204

Qy        275 VQPMALEQVMS------DRDSEDEV--DDDVADFEDRQMLDDFVDVNKDEK----QFMHL 322
              |  :| : |       || | | :   :   : : |:  ::    |: ||:     | |
Db        205 -PPKKIEFLQSCVLEFVDRTSSDLICGAEPYVEGQYRKYNNNSGFVSNDERNTPQSFSHF 263

Qy        323 WNSFVRKQRVIAD----GHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
                | :| |    |          : : ::| |  :       :  |   |
Db        264 TYEHSNHQLLIIDIQGVG---------DHYTDPQIHTYDGV----GFGIGNLGQKGFEKF 310

Qy        379 ATINNCNTILE--NCRNSSDTTTTNNNNSVDRPS---------DSNTNNNNIVDHPNDIN 427
              : || | :  | :: :  :  ::  :| ||          |:| |||| :: |: |
Db        311 LDTHKCNAICQYLNLQSINPKSEKSDCGTVPRPDLIFPDTSERDNNNNNNN--NNNNNNN 368

Qy        428 NKNNVDNKDNNSRDKVIK 445
              | ||  | :||:   : |
Db        369 NNNNNSNNNNNNNSSISK 386


RESULT 14
US-09-994-485-8
; Sequence 8, Application US/09994485
; Patent No. US20020142429A1
;   GENERAL INFORMATION:
;        APPLICANT: Ryazanov, Alexey G.
;                   Hait, William N.
;                   Pavur, Karen S.
;        TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
;                            AND METHODS OF USE THEREFOR
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: David A. Jackson, Esq.
;             STREET: 411 Hackensack Ave, Continental Plaza, 4th
;                     Floor
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/994,485
;             FILING DATE: 27-No. US20020142429A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/914,999
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 601-1-078
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-487-5800
;             TELEFAX: 201-343-1684
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 732 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Dictyostelium discoideum
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-994-485-8

  Query Match             4.4%;  Score 105;  DB 10;  Length 732;
  Best Local Similarity   20.1%;  Pred. No. 0.38;
  Matches   76;  Conservative   61;  Mismatches  167;  Indels   74;  Gaps   17;

Qy        117 EFKLFE-EYQTVNVSVKLNSFIFEEEG---SDDDKFEPFSLCSKPRKRRQRGGRNNTRRL 172
              | |:|   | |    :::   :    |   : | ||| |   :     : : |     |:
Db         34 EGKIFRFPYGTSQTDLQIGKMLPSGSGGGATADSKFEKFKARNTLADIQYKVGDTLYVRV 93

Qy        173 KV------CFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNI----PP 222
              |       ||  : :  :|:|  |     |       |  |    |:  |: :    |
Db         94 KKSKPTNDSLLPTLNIAFLDGSERAIKWEYD-------PYTT--TAQWTCTATLVKVEPV 144

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQF--------YHSHR 274
              |  :      :  |::       |:  : :  |:      | |          |:| :
Db        145 PFAEGAFRKAYHTLDLSKSGASGRYVSKIGKKPTPRPSYFEDVKMQMIAKKWADKYNSFK 204

Qy        275 VQPMALEQVMS------DRDSEDEV--DDDVADFEDRQMLDDFVDVNKDEK----QFMHL 322
              |  :| : |       || | | :   :   : : |:  ::    |: ||:     | |
Db        205 -PPKKIEFLQSCVLEFVDRTSSDLICGAEPYVEGQYRKYNNNSGFVSNDERNTPQSFSHF 263

Qy        323 WNSFVRKQRVIAD----GHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
                | :| |    |          : : ::| |  :       :  |   |
Db        264 TYEHSNHQLLIIDIQGVG---------DHYTDPQIHTYDGV----GFGIGNLGQKGFEKF 310

Qy        379 ATINNCNTILE--NCRNSSDTTTTNNNNSVDRPS---------DSNTNNNNIVDHPNDIN 427
              : || | :  | :: :  :  ::  :| ||          |:| |||| :: |: |
Db        311 LDTHKCNAICQYLNLQSINPKSEKSDCGTVPRPDLIFPDTSERDNNNNNNN--NNNNNNN 368

Qy        428 NKNNVDNKDNNSRDKVIK 445
              | ||  | :||:   : |
Db        369 NNNNNSNNNNNNNSSISK 386


RESULT 15
US-09-858-935B-77
; Sequence 77, Application US/09858935B
; Publication No. US20030069177A1
; GENERAL INFORMATION:
;  APPLICANT: Dubaquie, Yves
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Lowman, Henry B.
;  TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
;  FILE REFERENCE: P1794R1
;  CURRENT APPLICATION NUMBER: US/09/858,935B
;  CURRENT FILING DATE:  2002-07-02
;  PRIOR APPLICATION NUMBER: US 60/248,985
;  PRIOR FILING DATE: 2000-11-15
;  PRIOR APPLICATION NUMBER: US 60/204,490
;  PRIOR FILING DATE: 2000-05-16
;  NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 77
;   LENGTH: 97
;   TYPE: PRT
```

; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence is synthesized
US-09-858-935B-77

Query Match 4.4%; Score 104.5; DB 9; Length 97;
Best Local Similarity 34.4%; Pred. No. 0.022;
Matches 21; Conservative 16; Mismatches 23; Indels 1; Gaps 1;

```
Qy        380 TINNCNTILENCRNSSDTTTTNN-NNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNN 438
              |   |     |  ||:::  :|| ||| :  : :|:||:|  :: |: || || :| :|:
Db         13 TATGCANNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNS 72

Qy        439 S 439
              :
Db         73 N 73
```

Search completed: June 13, 2003, 15:38:45
Job time : 39 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:35:49 ; Search time 310 Seconds
(without alignments)
925.504 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2378	100.0	445	22	US-09-890-220-2	Sequence 2, Appli
2	2285	96.1	440	22	US-09-890-220-5	Sequence 5, Appli
3	1923	80.9	498	19	US-09-513-996A-69039	Sequence 69039, A
4	1796	75.5	367	19	US-09-513-996A-69040	Sequence 69040, A
5	1635	68.8	339	19	US-09-513-996A-69041	Sequence 69041, A
6	1153.5	48.5	428	26	US-10-219-999-37133	Sequence 37133, A
7	1153.5	48.5	428	27	US-60-324-109-23364	Sequence 23364, A
8	923	38.8	611	17	US-09-339-947A-1	Sequence 1, Appli
9	906.5	38.1	604	17	US-09-339-947A-8	Sequence 8, Appli
10	833	35.0	481	26	US-10-219-999-48099	Sequence 48099, A
11	833	35.0	481	27	US-60-312-544-10052	Sequence 10052, A
12	654	27.5	813	17	US-09-398-237-2	Sequence 2, Appli
13	549	23.1	107	22	US-09-890-220-8	Sequence 8, Appli
14	510.5	21.5	692	25	US-10-177-478-6	Sequence 6, Appli
15	451.5	19.0	186	22	US-09-890-220-14	Sequence 14, Appl
16	423	17.8	743	21	US-09-708-427-25830	Sequence 25830, A
17	423	17.8	765	21	US-09-708-427-25829	Sequence 25829, A
18	423	17.8	851	21	US-09-708-427-25828	Sequence 25828, A
19	368.5	15.5	108	22	US-09-890-220-12	Sequence 12, Appl
20	327	13.8	63	19	US-09-513-996A-69043	Sequence 69043, A
21	324	13.6	145	19	US-09-513-996A-79541	Sequence 79541, A
22	324	13.6	174	19	US-09-513-996A-79540	Sequence 79540, A
23	322.5	13.6	623	21	US-09-708-427-25832	Sequence 25832, A
24	311	13.1	154	22	US-09-890-220-10	Sequence 10, Appl
25	300	12.6	582	21	US-09-708-427-19473	Sequence 19473, A
26	266	11.2	295	21	US-09-708-427-25824	Sequence 25824, A
27	249	10.5	366	26	US-10-219-999-61010	Sequence 61010, A
28	249	10.5	366	27	US-60-324-109-29363	Sequence 29363, A
29	239	10.1	289	21	US-09-708-427-25825	Sequence 25825, A
30	225	9.5	388	1	PCT-US01-01341-931	Sequence 931, App
31	225	9.5	388	21	US-09-764-864-931	Sequence 931, App
32	225	9.5	388	24	US-10-080-129-931	Sequence 931, App
33	224.5	9.4	302	21	US-09-708-427-19475	Sequence 19475, A
34	224.5	9.4	427	21	US-09-708-427-19474	Sequence 19474, A
35	221	9.3	367	22	US-09-890-220-54	Sequence 54, Appl
36	221	9.3	739	1	PCT-US01-17936-5	Sequence 5, Appli
37	221	9.3	739	22	US-09-874-162A-5	Sequence 5, Appli
38	221	9.3	776	1	PCT-US01-17936-8	Sequence 8, Appli
39	221	9.3	776	22	US-09-874-162A-8	Sequence 8, Appli
40	221	9.3	803	22	US-09-890-220-16	Sequence 16, Appl
41	219	9.2	835	16	US-09-270-767-60845	Sequence 60845, A
42	219	9.2	835	16	US-09-270-849B-191180	Sequence 191180,
43	219	9.2	947	16	US-09-270-767-45347	Sequence 45347, A
44	219	9.2	947	27	US-60-167-324-1120	Sequence 1120, Ap
45	219	9.2	955	20	US-09-614-150-41697	Sequence 41697, A

ALIGNMENTS

RESULT 1
US-09-890-220-2
; Sequence 2, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: Gendall, Anthony
; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
; TITLE OF INVENTION: vernalisation gene VRN2.
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-890-220-2

Query Match 100.0%; Score 2378; DB 22; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.4e-235;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420

Qy        421 DHPNDINNKNNVDNKDNNSRDKVIK 445
              |||||||||||||||||||||||||
Db        421 DHPNDINNKNNVDNKDNNSRDKVIK 445


RESULT 2
US-09-890-220-5
; Sequence 5, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-5

  Query Match           96.1%;  Score 2285;  DB 22;  Length 440;
  Best Local Similarity  96.8%;  Pred. No. 5.6e-226;
  Matches  428;  Conservative    3;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLD 180

Qy        181 SPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSE 240
              |||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||:|
Db        181 SPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTE 240

Qy        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADF 300

Qy        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLF 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        301 EDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEVFSRFYEKELHCYSSLF 360

Qy        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV 420
              ||||||||||||||||||||||||||||||:|    | ||||||| ||||||||||||
Db        361 WCWRLFLIKLWNHGLVDSATINNCNTILENCRNTS--VTNNNNNSVDHPSDSNTNNNNIV 418

Qy        421 DHPNDINNKNNVDNKDNNSRDK 442
              |||||| ||||||||||||||
Db        419 DHPNDIKNKNNVDNKDNNSRDK 440


RESULT 3
US-09-513-996A-69039
; Sequence 69039, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69039
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..498
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..498 / Ceres Seq. ID 2185973
US-09-513-996A-69039

  Query Match           80.9%;  Score 1923;  DB 19;  Length 498;
  Best Local Similarity  89.4%;  Pred. No. 1.6e-188;
  Matches  370;  Conservative    3;  Mismatches    7;  Indels   34;  Gaps    3;

Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 167

Qy        103 LQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQ 162
              |||||||||||||||    | :|| ::   |                  ||||||||
Db        168 LQFHLNSSHDLFEFEFK----YISVNPTIMFCS------------------SKPRKRRQ 204

Qy        163 RGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPP 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 RGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPP 264

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEAR----------SHLLLQKRQFYH 271
              |||||||||||||||||||||||||||||||||||||          ||||||||||||
Db        265 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSTAILTEKISSSHLLLQKRQFYH 324

Qy        272 SHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQR 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 SHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQR 384

Qy        332 VIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENC 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 VIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENW 444

Qy        392 RNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 RNSSDTTTTNNNNSVDRPSDSNTNNNNIVSHPNDINNKNNVDNKDNNSRDKVIK 498
```

```
RESULT 4
US-09-513-996A-69040
; Sequence 69040, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69040
;   LENGTH: 367
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..367
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..367 / Ceres Seq. ID 2185974
US-09-513-996A-69040

  Query Match             75.5%;  Score 1796;  DB 19;  Length 367;
  Best Local Similarity   88.7%;  Pred. No. 1.3e-175;
  Matches  346;  Conservative    3;  Mismatches    7;  Indels   34;  Gaps    3;

Qy         67 MVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQT 126
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||    | :
Db          1 MVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFK----YIS 56

Qy        127 VNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
              || ::   |                   ||||||||||||||||||||||||||||||||
Db         57 VNPTIMFCS------------------SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN 97

Qy        187 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 GTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPAT 157

Qy        247 KTRKLSAERSEAR----------SHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDD 295
              |||||||||||||          ||||||||||||||||||||||||||||||||||||
Db        158 KTRKLSAERSEARSTAILTEKISSSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDD 217

Qy        296 DVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHR 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 DVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHR 277

Qy        356 YSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTN 415
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        278 YSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENWRNSSDTTTTNNNNSVDRPSDSNTN 337

Qy        416 NNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
              ||||| ||||||||||||||||||||||||
Db        338 NNNIVSHPNDINNKNNVDNKDNNSRDKVIK 367


RESULT 5
US-09-513-996A-69041
; Sequence 69041, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69041
;   LENGTH: 339
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 1..339
;   OTHER INFORMATION: any n or Xaa = unknown
;   FEATURE:
;   OTHER INFORMATION: Location 1..339 / Ceres Seq. ID 2185975
US-09-513-996A-69041

  Query Match             68.8%;  Score 1635;  DB 19;  Length 339;
  Best Local Similarity   87.8%;  Pred. No. 4.6e-159;
  Matches  318;  Conservative    3;  Mismatches    7;  Indels   34;  Gaps    3;

Qy         95 MLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLC 154
              |||||||||||||||||||||||||    | :|| ::   |
Db          1 MLCGSFKGLQFHLNSSHDLFEFEFK----YISVNPTIMFCS------------------- 37

Qy        155 SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQF 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 SKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQF 97

Qy        215 EMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEAR----------SHLL 263
              ||||||||||||||||||||||||||||||||||||||||||||||          ||||
Db         98 EMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSTAILTEKISSSHLL 157

Qy        264 LQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLW 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 LQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLW 217

Qy        324 NSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINN 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 NSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINN 277

Qy        384 CNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKV 443
              ||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        278 CNTILENWRNSSDTTTTNNNNSVDRPSDSNTNNNNIVSHPNDINNKNNVDNKDNNSRDKV 337

Qy        444 IK 445
              ||
Db        338 IK 339


RESULT 6
US-10-219-999-37133
; Sequence 37133, Application US/10219999
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)C
;  CURRENT APPLICATION NUMBER: US/10/219,999
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 60/324,109
;  PRIOR FILING DATE: 2001-09-21
;  PRIOR APPLICATION NUMBER: US 60/312,544
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 37133
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-219-999-37133

  Query Match             48.5%;  Score 1153.5;  DB 26;  Length 428;
  Best Local Similarity   54.0%;  Pred. No. 3.3e-109;
  Matches  235;  Conservative   57;  Mismatches  102;  Indels   41;  Gaps    8;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||||     : || |: ||:|||||||| |||| : |:| ||||| ||| ||| | |||
Db          1 MCRQNSPVHHAGEEEIAADESLLIYCKPVELYNILYRRALQNPSFLRRCLRYKIRASRKR 60
```

```
Qy          61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
               : |: |:|:|||:|  | |:|||| || ||||| | ||:|||||:||| |||||| ||| :
Db          61 RLRA-GIVIFNYRDRYNILRKTEVTEDFSCPFCLMQCGNFKGLRFHLCSSHDLFNFEFWV 119

Qy         121 FEEYQTVNVSVKLNSFIFEEEGSDD---DKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFL 177
                |:|| ||||||::  |   |   |    : : |  ||:||||||::      :|  | ||
Db         120 TEDYQAVNVSVKID--ILRSENVADGVIPQSQTFFFCSRPRKRRRKDSVQIEKRTNVKFL 177

Qy         178 PLDSPSLTNGTENGITLLND--------------------------GNRGLGYPEATEL 210
               ||||    |  ||    :|                        |  |  :|   :
Db         178 ELDSP---EGIHNGFLQKDDDILSCKGENVSRTSRSEKIFPSGRNDGGKFGPDHPGTMDN 234

Qy         211 AGQFEMTSNIPP---AIAHSSLDAG-AKVILTSEAVVPATKTRKLSAERSEARSHLLLQK 266
               | : |||    |:  ||:|   :| |  |:  :|| ||:||| :||::|: :||||
Db         235 LEHVESSFNIPGVSIAMPQSSVDPECSKSICKSDPALPA-KTKKLSMDRSDSRNRMLLQK 293

Qy         267 RQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSF 326
               | |:|||||||||||||:||||||||||||:|| |||:|||||||:||||| |||||||
Db         294 RLFFHSHRVQPMALEQVLSDRDSEDEVDDDIADLEDRRMLDDFVDVSKDEKQLMHLWNSF 353

Qy         327 VRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNT 386
               :|||||:||||: |||||||: : |||    :||||||||:|||||||||:|: |:|||:
Db         354 MRKQRVLADGHVPWACEAFSKLHGKELISSPALFWCWRLFMIKLWNHGLLDACTMNNCSI 413

Qy         387 ILENCRNSSDTTTTN 401
               :|:: ||    |  |
Db         414 VLDSYRNEGSGTRKN 428


RESULT 7
US-60-324-109-23364
; Sequence 23364, Application US/60324109
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)B
;  CURRENT APPLICATION NUMBER: US/60/324,109
;  CURRENT FILING DATE:  2001-09-21
;  NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 23364
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
US-60-324-109-23364

  Query Match             48.5%;  Score 1153.5;  DB 27;  Length 428;
  Best Local Similarity   54.0%;  Pred. No. 3.3e-109;
  Matches  235;  Conservative   57;  Mismatches  102;  Indels   41;  Gaps    8;

Qy           1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
               |||||     : || |: ||:|||||||| |||| : |:| |||||  ||| ||| | |||
Db           1 MCRQNSPVHHAGEEEIAADESLLIYCKPVELYNILYRRALQNPSFLRRCLRYKIRASRKR 60

Qy          61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
               : |: |:|:|||:|  | |:|||| || ||||| | ||:|||||:||| |||||| ||| :
Db          61 RLRA-GIVIFNYRDRYNILRKTEVTEDFSCPFCLMQCGNFKGLRFHLCSSHDLFNFEFWV 119

Qy         121 FEEYQTVNVSVKLNSFIFEEEGSDD---DKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFL 177
                |:|| ||||||::  |   |   |    : : |  ||:||||||::      :|  | ||
Db         120 TEDYQAVNVSVKID--ILRSENVADGVIPQSQTFFFCSRPRKRRRKDSVQIEKRTNVKFL 177

Qy         178 PLDSPSLTNGTENGITLLND--------------------------GNRGLGYPEATEL 210
               ||||    |  ||    :|                        |  |  :|   :
Db         178 ELDSP---EGIHNGFLQKDDDILSCKGENVSRTSRSEKIFPSGRNDGGKFGPDHPGTMDN 234

Qy         211 AGQFEMTSNIPP---AIAHSSLDAG-AKVILTSEAVVPATKTRKLSAERSEARSHLLLQK 266
               | : |||    |:  ||:|   :| |  |:  :|| ||:||| :||::|: :||||
Db         235 LEHVESSFNIPGVSIAMPQSSVDPECSKSICKSDPALPA-KTKKLSMDRSDSRNRMLLQK 293

Qy         267 RQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSF 326
               | |:|||||||||||||:||||||||||||:|| |||:|||||||:||||| |||||||
Db         294 RLFFHSHRVQPMALEQVLSDRDSEDEVDDDIADLEDRRMLDDFVDVSKDEKQLMHLWNSF 353

Qy         327 VRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNT 386
               :|||||:||||: |||||||: : |||    :||||||||:|||||||||:|: |:|||:
Db         354 MRKQRVLADGHVPWACEAFSKLHGKELISSPALFWCWRLFMIKLWNHGLLDACTMNNCSI 413

Qy         387 ILENCRNSSDTTTTN 401
               :|:: ||    |  |
Db         414 VLDSYRNEGSGTRKN 428


RESULT 8
US-09-339-947A-1
; Sequence 1, Application US/09339947A
; GENERAL INFORMATION:
;  APPLICANT:  YOSHIDA, Nobumasa
;  APPLICANT:  YANAI, Yukihiro
;  APPLICANT:  KATO, Yoshihiro
;  APPLICANT:  HIRATSUKA, Junzo
;  APPLICANT:  TAKAHASHI, Shigeru
;  APPLICANT:  MIWA, Tatsushi
;  TITLE OF INVENTION:  GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
;  TITLE OF INVENTION:  FLOWERING
;  FILE REFERENCE:  032735-001
;  CURRENT APPLICATION NUMBER: US/09/339,947A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:  JP 1998-180065
;  PRIOR FILING DATE:  1998-06-26
;  PRIOR APPLICATION NUMBER:  JP 1999-179043
;  PRIOR FILING DATE:  1999-06-24
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 611
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: ZN_FING
;   LOCATION: (306)..(327)
US-09-339-947A-1

  Query Match             38.8%;  Score 923;  DB 17;  Length 611;
  Best Local Similarity   34.8%;  Pred. No. 3.6e-85;
  Matches  216;  Conservative   52;  Mismatches  108;  Indels  244;  Gaps    9;

Qy           1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
               || :: | : | || |: :|:|  ||||| ||||   |:: || || |||:||| || ||
Db           1 MCHEDSRLRISEEEEIAAEESLAAYCKPVELYNIIQRRAIRNPLFLQRCLHYKIEAKHKR 60

Qy          61 ------------------------------------------------------------ 60

Db          61 RIQMTVFLSGAIDAGVQTQKLFPLYILLARLVSPKPVAEYSAVYRFSRACILTGGLGVDG 120

Qy          61 ------------------------------------------------------------ 60

Db         121 VSQAQANFLLPDMNRLALEAKSGSLAILFISFAGAQNSQFGIDSGKIHSGNIGGHCLWSK 180

Qy          61 -------------------------------------KSRS-------------------- 64
                                                    || |
Db         181 IPLQSLYASWQKSPNMDLGQRVDTVSLVEMQPCFIKLKSMSEEKCVSIQVPSNPLTSSSP 240

Qy          65 ----------------------------------------TGMVVFNYKDCNNTLQKTEV 84
                                                       || |||||:  || ||||||
Db         241 QQVQVTISAEEVGSTEKSPYSSFSYNDISSSSLLQIIRLRTGNVVFNYRYYNNKLQKTEV 300

Qy          85 REDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSD 144
               || ||||| : | ||||||::|| |:|||  ||| : ||:| ||||:|  : | :    |
Db         301 TEDFSCPFCLVKCASFKGLRYHLPSTHDLLNFEFWVTEEFQAVNVSLKTETMISKVNEDD 360
```

```
Qy        145 -DDKFEPFSLCSKP-RKRRQRGGRNNTRRLKVCFLPLDSPSLTNG------TENGITLLN 196
Db        361 VDPKQQTFFFSSKKFRRRRQKSQVRSSRQ---------GPHLGLGCEVLDKTDDAHSVRS 411

Qy        197 DGNR---GLGYPEATELAGQFEMTSNIP----PAIAHSSLDAGAKVILTSEAVVPATKTR 249
Db        412 EKSRIPPGKHY----ERIGGAESGQRVPPGTSPADVQSCGDPDYVQSIAGSTMLQFAKTR 467

Qy        250 KLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDF 309
Db        468 KISIERSDLRNRSLLQKRQFFHSHRAQPMALEQVLSDRDSEDEVDDDVADFEDRRMLDDF 527

Qy        310 VDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIK 369
Db        528 VDVTKDEKQMMHMWNSFVRKQRVLADGHIPWACEAFSRLHGPIMVRTPHLIWCWRVFMVK 587

Qy        370 LWNHGLVDSATINNCNTILE 389
Db        588 LWNHGLLDARTMNNCNTFLE 607


RESULT 9
US-09-339-947A-8
; Sequence 8, Application US/09339947A
; GENERAL INFORMATION:
;  APPLICANT:  YOSHIDA, Nobumasa
;  APPLICANT:  YANAI, Yukihiro
;  APPLICANT:  KATO, Yoshihiro
;  APPLICANT:  HIRATSUKA, Junzo
;  APPLICANT:  TAKAHASHI, Shigeru
;  APPLICANT:  MIWA, Tatsushi
;  TITLE OF INVENTION:  GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
;  TITLE OF INVENTION:  FLOWERING
;  FILE REFERENCE:  032735-001
;  CURRENT APPLICATION NUMBER: US/09/339,947A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:  JP 1998-180065
;  PRIOR FILING DATE:  1998-06-26
;  PRIOR APPLICATION NUMBER:  JP 1999-179043
;  PRIOR FILING DATE:  1999-06-24
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 604
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: ZN_FING
;   LOCATION: (310)..(335)
US-09-339-947A-8

  Query Match           38.1%;  Score 906.5;  DB 17;  Length 604;
  Best Local Similarity   33.3%;  Pred. No. 1.8e-83;
  Matches  209;  Conservative   59;  Mismatches  109;  Indels  251;  Gaps    7;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
Db          1 MCRHQPRARLSPDEQLAAEESFALYCKPVELYNIIQRRSIKNPAFLQRCLLYKIHARRKK 60

Qy         61 KSRST---------------------------------------------------------- 65
Db         61 RSLITISLSGGTNKELRAQNIFPLYVLLARPTNNVSLEGHSPIYRFSRACLLTSFHEFGN 120

Qy         66 ---------------------------------------------------------------- 65

Db        121 KDYTEATFVIPDVKNLATSRACSLNIILISCGRAEQTFDDNNCSGNHVEGSTLQKLEGKC 180

Qy         66 ---------------------------------------------------------------- 65

Db        181 FWGKIPIDLLASSLGNCVSLSLGHTVEMSSTVEMTPSFLEPKFLEDDSCLTFCSQKVDAT 240

Qy         66 -----------------------------------------------GMVVFNYKDCNNTLQ 80

Db        241 GSFQLQVSISAQEAGAKDMSESPYSVYSYNDVPPSSLTHIIRLRSGNVLFNYKYYNNTMQ 300

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFE- 139
Db        301 KTEVTEDFSCPFCLVPCGSFKGLGCHLNASHDLFHYEFWISEECQAVNVSLKTDSWRTEL 360

Qy        140 -EEGSDDDKFEPFSLCSKPRKR-RQRGGRNNTRRLKVCFLPLDSP-SLTNGTENGITLLN 196
Db        361 LAEGV-DPRHQTFSYRSRFKKRKRVEISSDKIRHVHPHIVDSGSPEDAQAGSEDDYVQRE 419

Qy        197 DGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKV---ILTSEAVVPATKTRKLSA 253
Db        420 NGS----------------------SVAHASVDPANSLHGSNLSAPTVLQFGKTRKLSV 456

Qy        254 ERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVN 313
Db        457 ERADPRNRQLLQKRQFFHSHRAQPMAWSKVFSDRDSEDEVDDDIADFEDRRMLDDFVDVT 516

Qy        314 KDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNH 373
Db        517 KDEKLIMHMWNSFVRKQRVLADGHIPWACEAFSQFHGQELVQNPALLWCWRFFMVKLWNH 576

Qy        374 GLVDSATINNCNTILENCRNSSDTTTTN 401
Db        577 SLLDARAMNACNTILEGYLNGSSDPKKN 604


RESULT 10
US-10-219-999-48099
; Sequence 48099, Application US/10219999
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)C
;  CURRENT APPLICATION NUMBER: US/10/219,999
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 60/324,109
;  PRIOR FILING DATE: 2001-09-21
;  PRIOR APPLICATION NUMBER: US 60/312,544
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 48099
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Zea mays
US-10-219-999-48099

  Query Match           35.0%;  Score 833;  DB 26;  Length 481;
  Best Local Similarity   48.6%;  Pred. No. 4.9e-76;
  Matches  180;  Conservative   45;  Mismatches   83;  Indels   62;  Gaps    8;

Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 102
Db        153 PSLLPEIIRLR-----------AGNVLFNYKYYNNTMQKTEVTEDFSCPFCLVPCGSFKG 201

Qy        103 LQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF--IFEEEGSDDDKFEPFSLCSKPRKR 160
Db        202 LGCHLNSSHDLFHYELWISEECQAVNVSLKADAWKTEFVAEGV-DPRHQTFSYCSRFKKR 260

Qy        161 RQRGGRNNTRRLKVCFLPLDSPSLTNGT---------------ENGITLLNDGNRGLGYP 205
Db        261 R----RLETTAEK--FRHVHPHIMESGSHEDAQAGSEDDYVQRENGLSVAN--------- 305

Qy        206 EATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQ 265
Db        306 -----------TSVDPSQPVHGS-------NLSPPTVLQFGKSRKLSAERSDPRNRQLLQ 347

Qy        266 KRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNS 325
```

```
                |||||:|||| ||| ||||:|| |||||||||:||||||:|||||||| ||||  ||:|||
Db        348 KRQFFHSHRAQPMPLEQVLSDHDSEDEVDDDIADFEDRRMLDDFVDVTKDEKLIMHMWNS 407

Qy        326 FVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCN 385
              |||||||:||||| |||||||| : ::| :  :| | || |:|||||| |:|:  :| ||
Db        408 FVRKQRVLADGHIPWACEAFSRCHGQQLVQNPALLWGWRFFMIKLWNHSLLDARAMNACN 467

Qy        386 TILENCRNSS 395
              |||:  :: |
Db        468 TILQGYQDGS 477


RESULT 11
US-60-312-544-10052
; Sequence 10052, Application US/60312544
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Edgerton, Michael D
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Stein, Joshua
;  TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-10(52726)A
;  CURRENT APPLICATION NUMBER: US/60/312,544
;  CURRENT FILING DATE:  2001-08-15
;  NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 10052
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3206-226-G12_FLI
US-60-312-544-10052

  Query Match             35.0%;  Score 833;  DB 27;  Length 481;
  Best Local Similarity   48.6%;  Pred. No. 4.9e-76;
  Matches  180;  Conservative   45;  Mismatches   83;  Indels   62;  Gaps    8;

Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 102
              ||  ||   :  :             | |:||||  |||:||||| || ||||| : ||||||
Db        153 PSLLPEIIRLR-----------AGNVLFNYKYYNNTMQKTEVTEDFSCPFCLVPCGSFKG 201

Qy        103 LQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF--IFEEEGSDDDKFEPFSLCSKPRKR 160
              |  |||||||||| :|  : || | ||||:| :::   |  ||  | : : || ||: :||
Db        202 LGCHLNSSHDLFHYELWISEECQAVNVSLKADAWKTEFVAEGV-DPRHQTFSYCSRFKKR 260

Qy        161 RQRGGRNNTRRLKVCFLPLDSPSLTNGT--------------ENGITLLNDGNRGLGYP 205
              |    |  |   |  |  :   : :|:              |||::: |
Db        261 R----RLETTAEK--FRHVHPHIMESGSHEDAQAGSEDDYVQRENGLSVAN--------- 305

Qy        206 EATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQ 265
                        ||  |    | |         |:   |:   |:|||||||: |:  |||
Db        306 ----------TSVDPSQPVHGS-------NLSPPTVLQFGKSRKLSAERSDPRNRQLLQ 347

Qy        266 KRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNS 325
              ||||:|||| ||| ||||:|| |||||||||:||||||:|||||||| ||||  ||:|||
Db        348 KRQFFHSHRAQPMPLEQVLSDHDSEDEVDDDIADFEDRRMLDDFVDVTKDEKLIMHMWNS 407

Qy        326 FVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCN 385
              |||||||:||||| |||||||| : ::| :  :| | || |:|||||| |:|:  :| ||
Db        408 FVRKQRVLADGHIPWACEAFSRCHGQQLVQNPALLWGWRFFMIKLWNHSLLDARAMNACN 467

Qy        386 TILENCRNSS 395
              |||:  :: |
Db        468 TILQGYQDGS 477


RESULT 12
US-09-398-237-2
; Sequence 2, Application US/09398237
; GENERAL INFORMATION:
;  APPLICANT: Commonwealth Scientific and Industrial Research Organisation
;  TITLE OF INVENTION: Novel method of regulating seed development in plants and
;  TITLE OF INVENTION: genetic sequences therefor
;  FILE REFERENCE: p:\oper\mro\autseed.pct
;  CURRENT APPLICATION NUMBER: US/09/398,237
;  CURRENT FILING DATE: 1999-09-20
;  EARLIER APPLICATION NUMBER: US 60/101184
;  EARLIER FILING DATE: 1998-09-21
;  EARLIER APPLICATION NUMBER: AU PP6063
;  EARLIER FILING DATE: 1998-09-22
;  EARLIER APPLICATION NUMBER: AU PP6062
;  EARLIER FILING DATE: 1998-09-22
;  EARLIER APPLICATION NUMBER: AU PP6061
;  EARLIER FILING DATE: 1998-09-22
;  EARLIER APPLICATION NUMBER: AU PQ1345
;  EARLIER FILING DATE: 1999-07-01
;  EARLIER APPLICATION NUMBER: AU PQ1346
;  EARLIER FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 213
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 813
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-398-237-2

  Query Match             27.5%;  Score 654;  DB 17;  Length 813;
  Best Local Similarity   24.2%;  Pred. No. 3.3e-57;
  Matches  184;  Conservative   65;  Mismatches  147;  Indels  364;  Gaps    9;

Qy         13 EEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRKSRSTGMVVFNY 72
              :: :  |||:: | ||| :|   ||  || || | | | | ||:|:|| | | : |||
Db         56 DDDVDVDENIIKYIKPVAVYKKLETRSKNNPYFLRRSLKYIIQAKKKKKSNSGGKIRFNY 115

Qy         73 KDCNNTLQ-KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNV 129
                :| :| :  | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|
Db        116 RDVSNKMTLKAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDV 175

Qy        130 SVKLNSFIFE--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
              ||| ::  |   ::   : :  | : ||| | :||||   || ::| |  : ||   |
Db        176 SVKSDTIKFGVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPR 235

Qy        187 GTENGITLLNDGN---------------------------------------------- 199
              ||||  | :|| |
Db        236 GTENDSTHVNDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAH 295

Qy        200 ---------------------------------------------------------- 199

Db        296 SSAEKNESTHVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDED 355

Qy        200 ---------------------------------RG------------------------ 201
                                                ||
Db        356 VSFPPRTRSSKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSE 415

Qy        202 -----------LGYPEATELAGQFEMTSNIPPAIAHSS---------------------- 228
                          |||  |    :: :   |||| ||
Db        416 PKVLHVNDENVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSK 475

Qy        229 ---------------------------------------------------------- 228

Db        476 KNKSTRKNVDNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSS 535

Qy        229 ----------------------------LDAGAKVILTSEAVVP---------------- 244
                                           |      |  ||  ||
Db        536 KKNKSTRKNDDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSS 595

Qy        245 ---------------------------------------------ATKTRKLSAERSEA 258
                                                           :: ::|.||| ||
Db        596 KKNKSTHKKDDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEA 655

Qy        259 RSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQ 318
              :    |: |||||| :|||  |||||: |||:| ||    |  :|  |:  | |:|:||:
Db        656 KRLERLKGRQFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKR 715
```

```
Qy        319 FMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
              :|:||| |||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: :
Db        716 YMYLWNIFVRKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICA 775

Qy        379 ATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
              | : | |||           :|:: :    | |  | ||
Db        776 KTFHKCTTIL----------LSNSDEAGQFTSGSAANANN 805


RESULT 13
US-09-890-220-8
; Sequence 8, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 107
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-8

  Query Match             23.1%;  Score 549;  DB 22;  Length 107;
  Best Local Similarity   99.0%;  Pred. No. 9.5e-48;
  Matches  100;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK 101
              ||||||||||||||||||:|||||||||||||||||||||
Db         61 KSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFK 101


RESULT 14
US-10-177-478-6
; Sequence 6, Application US/10177478
; GENERAL INFORMATION:
;  APPLICANT: Dang, Van-Dinh
;  APPLICANT:  Okamuro, Jack
;  TITLE OF INVENTION: Chimeric Histone Acetyltransferase
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 11696-037001
;  CURRENT APPLICATION NUMBER: US/10/177,478
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: U.S. 60/300,135
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 692
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-177-478-6

  Query Match             21.5%;  Score 510.5;  DB 25;  Length 692;
  Best Local Similarity   22.3%;  Pred. No. 1.7e-42;
  Matches  154;  Conservative   53;  Mismatches  121;  Indels  363;  Gaps    8;

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVSVKLNSFIF 138
              | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|||| ::  |
Db          4 KAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDVSVKSDTIKF 63

Qy        139 E--EEGSDDDKFEPPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195
                ::   : :  | : ||| | :||||   || ::| |  : ||   |  |||| | :
Db         64 GVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPRGTENDSTHV 123

Qy        196 NDGN---------------------------------------------------------- 199
              || |
Db        124 NDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAHSSAEKNEST 183

Qy        200 ---------------------------------------------------------- 199

Db        184 HVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDEDVSFPPRTRS 243

Qy        200 ------------------------RG---------------------------------- 201
                                       ||
Db        244 SKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSEPKVLHVNDE 303

Qy        202 --LGYPEATELAGQFEMTSNIPPAIAHSS------------------------------ 228
                  |||  |    :: :   |||| ||
Db        304 NVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSKKNKSTRKNV 363

Qy        229 ---------------------------------------------------------- 228

Db        364 DNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSSKKNKSTRKN 423

Qy        229 -------------------LDAGAKVILTSEAVVP------------------------ 244
                                |    |  ||  ||
Db        424 DDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSSKKNKSTHKK 483

Qy        245 -----------------------------------ATKTRKLSAERSEARSHLLLQKR 267
                                                  :: ::| ||| ||:    |: |
Db        484 DDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEAKRLERLKGR 543

Qy        268 QFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
              |||||  :|||  |||||: |||:| ||   |  :|  |:  |.|:|:||::|:||| ||
Db        544 QFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKRYMYLWNIFV 603

Qy        328 RKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTI 387
              ||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: : | : | ||
Db        604 RKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICAKTFHKCTTI 663

Qy        388 LENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
              |           :|:: :    | |  | ||
Db        664 L----------LSNSDEAGQFTSGSAANANN 684


RESULT 15
US-09-890-220-14
; Sequence 14, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics us
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 186
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-14

  Query Match             19.0%;  Score 451.5;  DB 22;  Length 186;
  Best Local Similarity   50.8%;  Pred. No. 2.6e-37;
  Matches   99;  Conservative   32;  Mismatches   43;  Indels   21;  Gaps    7;
```

```
Qy        237 LTSEAVVPATKTRKLSAERSEARSHLL-LQKRQFYHSHRVQPMALEQVMSDRDSEDEVD- 294
              ||:|| ||        |:||:| || | | |||||||    ||::||||||||||||::||
Db          6 LTTEAKVP--------AKRSKATSHYLPLHKRQFYHSRTGQPLSLEQVMSDRDSENDVDK 57

Qy        295 -DDVADFEDRQMLDDFVDVNK-DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKE 352
               || |  |: |||:  :| |:   ::|: ||||||::||::|| || |||||||| : :|
Db         58 NDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAHIPWACEAFSRLHLQE 117

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDS 412
              |    ||  ||| |:|| |::||:|  |:| ||||: :  | | |    |||:
Db        118 LRSNLSLDLCWRQFMIKQWDYGLLDRVTMNKCNTIIYH--NISTTNDDINNNNT-----R 170

Qy        413 NTNNNNIVDHPNDIN 427
              |:| ::||  :|||
Db        171 TTDNMDVVD--DDIN 183


Search completed: June 13, 2003, 15:43:51
Job time : 313 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:30:54 ; Search time 54 Seconds
(without alignments)
2102.183 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155225 seqs, 255096386 residues

Total number of hits satisfying chosen parameters: 1155225

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1153.5	48.5	428	6	US-10-424-599-254264	Sequence 254264,
2	1153.5	48.5	435	6	US-10-425-114-46823	Sequence 46823, A
3	833	35.0	481	6	US-10-425-114-54921	Sequence 54921, A
4	654	27.5	813	6	US-10-231-778-2	Sequence 2, Appli
5	648.5	27.3	640	6	US-10-437-963-193525	Sequence 193525,
6	617.5	26.0	201	6	US-10-437-963-114092	Sequence 114092,
7	575.5	24.2	242	6	US-10-424-599-283985	Sequence 283985,
8	510.5	21.5	692	6	US-10-198-723-6	Sequence 6, Appli
9	491	20.6	166	6	US-10-424-599-254263	Sequence 254263,
10	269	11.3	102	6	US-10-424-599-183487	Sequence 183487,
11	256.5	10.8	116	6	US-10-424-599-216298	Sequence 216298,
12	255.5	10.7	151	6	US-10-437-963-199908	Sequence 199908,
13	249	10.5	394	6	US-10-425-114-70148	Sequence 70148, A
14	190	8.0	263	6	US-10-437-963-184120	Sequence 184120,
15	159	6.7	195	6	US-10-424-599-250979	Sequence 250979,
16	140.5	5.9	116	6	US-10-424-599-152510	Sequence 152510,
17	132	5.6	749	6	US-10-369-493-21933	Sequence 21933, A
18	128	5.4	106	6	US-10-437-963-116154	Sequence 116154,
19	124	5.2	1288	1	PCT-US02-38360-2	Sequence 2, Appli
20	122	5.1	86	6	US-10-437-963-201278	Sequence 201278,
21	122	5.1	123	6	US-10-437-963-109658	Sequence 109658,
22	119	5.0	1094	1	PCT-US01-29288A-216	Sequence 216, App
23	119	5.0	1094	6	US-10-149-310-216	Sequence 216, App
24	118	5.0	249	6	US-10-437-963-135834	Sequence 135834,
25	115.5	4.9	2184	6	US-10-304-095-6	Sequence 6, Appli
26	114.5	4.8	879	6	US-10-282-122A-52571	Sequence 52571, A
27	114	4.8	964	1	PCT-US01-29288A-256	Sequence 256, App
28	114	4.8	964	6	US-10-149-310-256	Sequence 256, App
29	113.5	4.8	105	6	US-10-437-963-112367	Sequence 112367,
30	113.5	4.8	654	5	US-09-134-000C-6645	Sequence 6645, Ap
31	113.5	4.8	654	5	US-09-134-000C-6645	Sequence 6645, Ap
32	113.5	4.8	654	6	US-10-434-665-6645	Sequence 6645, Ap
33	112	4.7	104	6	US-10-437-963-198200	Sequence 198200,
34	108.5	4.6	905	6	US-10-369-493-21890	Sequence 21890, A
35	108	4.5	336	6	US-10-424-599-233789	Sequence 233789,
36	108	4.5	336	6	US-10-438-246-32035	Sequence 32035, A
37	108	4.5	337	6	US-10-425-114-39945	Sequence 39945, A
38	107.5	4.5	192	6	US-10-437-963-152724	Sequence 152724,
39	107.5	4.5	833	6	US-10-369-493-21936	Sequence 21936, A
40	107	4.5	138	6	US-10-437-963-176416	Sequence 176416,
41	107	4.5	313	6	US-10-425-114-52231	Sequence 52231, A
42	106.5	4.5	833	6	US-10-310-154-480	Sequence 480, App
43	106	4.5	328	6	US-10-437-963-107965	Sequence 107965,
44	105.5	4.4	787	6	US-10-282-122A-72461	Sequence 72461, A
45	104.5	4.4	97	6	US-10-271-869-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1
US-10-424-599-254264
; Sequence 254264, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254264
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1.pep
US-10-424-599-254264

Query Match 48.5%; Score 1153.5; DB 6; Length 428;
Best Local Similarity 54.0%; Pred. No. 3.9e-100;
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;

```
Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||||     : || |: ||:|||||||| |||| : |:| ||||| ||| ||| | |||
Db          1 MCRQNSPVHHAGEEEIAADESLLIYCKPVELYNILYRRALQNPSFLRRCLRYKIRASRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              : |: |:|:|||:|  | |:|||| || ||||| | ||:||||:||| |||||| ||| :
Db         61 RLRA-GIVIFNYRDRYNILRKTEVTEDFSCPFCLMQCGNFKGLRFHLCSSHDLFNFEFWV 119

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDD---DKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFL 177
                 |:|| ||||||::  |   |   |    : : |  ||:||||||::      :|  | ||
Db        120 TEDYQAVNVSVKID--ILRSENVADGVIPQSQTFFFCSRPRKRRRKDSVQIEKRTNVKFL 177

Qy        178 PLDSPSLTNGTENGITLLND-------------------------GNRGLGYPEATEL 210
               ||||   |  ||   :|                         |  |  :|  :
Db        178 ELDSP---EGIHNGFLQKDDDILSCKGENVSRTSRSEKIFPSGRNDGGKFGPDHPGTMDN 234

Qy        211 AGQFEMTSNIPP---AIAHSSLDAG-AKVILTSEAVVPATKTRKLSAERSEARSHLLLQK 266
                 | : |||    |:  ||:|   :| |  |:  :|| ||:||| :||::|: :||||
Db        235 LEHVESSFNIPGVSIAMPQSSVDPECSKSICKSDPALPA-KTKKLSMDRSDSRNRMLLQK 293

Qy        267 RQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSF 326
              | |:|||||||||||||:|||||||||||:|| |||:|||||||:||||| |||||||
Db        294 RLFFHSHRVQPMALEQVLSDRDSEDEVDDDIADLEDRRMLDDFVDVSKDEKQLMHLWNSF 353
```

```
Qy        327 VRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNT 386
              :|||||:||||: ||||||| : |||    :|||||||:||||||||:|: |:|||:
Db        354 MRKQRVLADGHVPWACEAFSKLHGKELISSPALFWCWRLFMIKLWNHGLLDACTMNNCSI 413

Qy        387 ILENCRNSSDTTTTN 401
              :|:: ||    |  |
Db        414 VLDSYRNEGSGTRKN 428


RESULT 2
US-10-425-114-46823
; Sequence 46823, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46823
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 701055296_FLI.pep
US-10-425-114-46823

  Query Match             48.5%;  Score 1153.5;  DB 6;  Length 435;
  Best Local Similarity   54.0%;  Pred. No. 4e-100;
  Matches  235;  Conservative   57;  Mismatches  102;  Indels   41;  Gaps    8;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||||     : || |: ||:|||||||| |||| : |:| ||||| ||| ||| | |||
Db          8 MCRQNSPVHHAGEEEIAADESLLIYCKPVELYNILYRRALQNPSFLRRCLRYKIRASRKR 67

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              : |: |:|:|||:|  | |:|||| || ||||| | ||:||||:||| |||||| ||| :
Db         68 RLRA-GIVIFNYRDRYNILRKTEVTEDFSCPFCLMQCGNFKGLRFHLCSSHDLFNFEFWV 126

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDD---DKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFL 177
               |:|| ||||||:: |  |  |    : : |  ||:||||||::       :|  | ||
Db        127 TEDYQAVNVSVKID--ILRSENVADGVIPQSQTFFFCSRPRKRRRKDSVQIEKRTNVKFL 184

Qy        178 PLDSPSLTNGTENGITLLND-------------------------GNRGLGYPEATEL 210
              .||||    |  ||    :|                           |  |  :|   :
Db        185 ELDSP---EGIHNGFLQKDDDILSCKGENVSRTSRSEKIFPSGRNDGGKFGPDHPGTMDN 241

Qy        211 AGQFEMTSNIPP---AIAHSSLDAG-AKVILTSEAVVPATKTRKLSAERSEARSHLLLQK 266
                  | : |||    |:  ||:|   :|.|  |:  :|| ||:||| :||::|:  :||||
Db        242 LEHVESSFNIPGVSIAMPQSSVDPECSKSICKSDPALPA-KTKKLSMDRSDSRNRMLLQK 300

Qy        267 RQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSF 326
              | |:|||||||||||||:||||||||||||:|| |||:||||||||:||||| |||||||
Db        301 RLFFHSHRVQPMALEQVLSDRDSEDEVDDDIADLEDRRMLDDFVDVSKDEKQLMHLWNSF 360

Qy        327 VRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNT 386
              :|||||:||||: ||||||| : |||    :|||||||:||||||||:|: |:|||:
Db        361 MRKQRVLADGHVPWACEAFSKLHGKELISSPALFWCWRLFMIKLWNHGLLDACTMNNCSI 420

Qy        387 ILENCRNSSDTTTTN 401
              :|:: ||    |  |
Db        421 VLDSYRNEGSGTRKN 435


RESULT 3
US-10-425-114-54921
; Sequence 54921, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54921
;   LENGTH: 481
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12_FLI.pep
US-10-425-114-54921

  Query Match             35.0%;  Score 833;  DB 6;  Length 481;
  Best Local Similarity   48.6%;  Pred. No. 8.6e-70;
  Matches  180;  Conservative   45;  Mismatches   83;  Indels   62;  Gaps    8;

Qy         43 PSFLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG 102
              || ||  :  :             | |:||||  |||:||||| || ||||| : ||||||
Db        153 PSLLPEIIRLR-----------AGNVLFNYKYYNNTMQKTEVTEDFSCPFCLVPCGSFKG 201

Qy        103 LQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSF--IFEEEGSDDDKFEPFSLCSKPRKR 160
              |  |||||||||| :|  : || | ||||:| :::   |  ||  | : : || ||: :||
Db        202 LGCHLNSSHDLFHYELWISEECQAVNVSLKADAWKTEFVAEGV-DPRHQTFSYCSRFKKR 260

Qy        161 RQRGGRNNTRRLKVCFLPLDSPSLTNGT---------------ENGITLLNDGNRGLGYP 205
              |    |  |   |  |  :    : :|:                |||::: |
Db        261 R----RLETTAEK--FRHVHPHIMESGSHEDAQAGSEDDYVQRENGLSVAN--------- 305

Qy        206 EATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQ 265
                        ||  |   | |   |:  :|   |:|||||||||: |:  |||
Db        306 -----------TSVDPSQPVHGS-------NLSPPTVLQFGKSRKLSAERSDPRNRQLLQ 347

Qy        266 KRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNS 325
              ||||:|||| ||| ||||:|| |||||||||:||||||:|||||||| ||||  ||:|||
Db        348 KRQFFHSHRAQPMPLEQVLSDHDSEDEVDDDIADFEDRRMLDDFVDVTKDEKLIMHMWNS 407

Qy        326 FVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCN 385
              |||||||:||||| |||||||| : ::| :  :| | || |:|||||| |:|:  :| ||
Db        408 FVRKQRVLADGHIPWACEAFSRCHGQQLVQNPALLWGWRFFMIKLWNHSLLDARAMNACN 467

Qy        386 TILENCRNSS 395
              |||:  :: |
Db        468 TILQGYQDGS 477


RESULT 4
US-10-231-778-2
; Sequence 2, Application US/10231778
; GENERAL INFORMATION:
;  APPLICANT: Bilodeau, Pierre
;  APPLICANT:  Chaudhury, Abdul M.
;  APPLICANT:  Dennis, Elizabeth S.
;  APPLICANT:  Koltunow, Anna M.G.
;  APPLICANT:  Luo, Ming
;  APPLICANT:  Peacock, William J.
;  TITLE OF INVENTION: Method for inducing seed development by down-regulating
;  TITLE OF INVENTION:  expression of the FIS2 gene
;  FILE REFERENCE: 72-98A
;  CURRENT APPLICATION NUMBER: US/10/231,778
;  CURRENT FILING DATE:  2002-11-08
;  PRIOR APPLICATION NUMBER: 09/398,237
;  PRIOR FILING DATE: 1999-09-20
;  PRIOR APPLICATION NUMBER: 60/101,184
```

```
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: AU PP6061
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6062
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6063
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PQ1345
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: AU PQ1346
;  PRIOR FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 239
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 813
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-231-778-2

  Query Match             27.5%;  Score 654;  DB 6;  Length 813;
  Best Local Similarity   24.2%;  Pred. No. 1.4e-52;
  Matches  184;  Conservative   65;  Mismatches  147;  Indels  364;  Gaps    9;

Qy         13 EEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRKSRSTGMVVFNY 72
              :: :  |||:: | ||| :|     ||  || || | | | ||:|:|| | | : |||
Db         56 DDDVDVDENIIKYIKPVAVYKKLETRSKNNPYFLRRSLKYIIQAKKKKKSNSGGKIRFNY 115

Qy         73 KDCNNTLQ-KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNV 129
              :| :| :  | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|
Db        116 RDVSNKMTLKAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDV 175

Qy        130 SVKLNSFIFE--EEGSDDDKFEPPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTN 186
              ||| ::  |   ::   : :  | : ||| | :||||   || ::| |  : ||   |
Db        176 SVKSDTIKFGVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPR 235

Qy        187 GTENGITLLNDGN---------------------------------------------- 199
              ||||  | :|| |
Db        236 GTENDSTHVNDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAH 295

Qy        200 ---------------------------------------------------------- 199

Db        296 SSAEKNESTHVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDED 355

Qy        200 ------------------------------RG------------------------- 201
                                           ||
Db        356 VSFPPRTRSSKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSE 415

Qy        202 -----------LGYPEATELAGQFEMTSNIPPAIAHSS---------------------- 228
                        |||  |    :: :   |||| ||
Db        416 PKVLHVNDENVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSK 475

Qy        229 ---------------------------------------------------------- 228

Db        476 KNKSTRKNVDNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSS 535

Qy        229 ---------------------------LDAGAKVILTSEAVVP---------------- 244
                                         |      |  ||  ||
Db        536 KKNKSTRKNDDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSS 595

Qy        245 ---------------------------------------------ATKTRKLSAERSEA 258
                                                          :: ::| ||| ||
Db        596 KKNKSTHKKDDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEA 655

Qy        259 RSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQ 318
              :    |: ||||||  :|||  |||||: |||:| ||   |  :|  |:  | |:|:||:
Db        656 KRLERLKGRQFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKR 715

Qy        319 FMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
              :|:||| |||||||||||:  |||| |:: :::|:   ||  | ||:| |||||:||: :
Db        716 YMYLWNIFVRKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICA 775

Qy        379 ATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
              | : | |||           :|:: :    | |  | ||

Db        776 KTFHKCTTIL----------LSNSDEAGQFTSGSAANANN 805


RESULT 5
US-10-437-963-193525
; Sequence 193525, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 193525
;   LENGTH: 640
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_89656C.1.pep
US-10-437-963-193525

  Query Match             27.3%;  Score 648.5;  DB 6;  Length 640;
  Best Local Similarity   25.5%;  Pred. No. 3.4e-52;
  Matches  168;  Conservative   74;  Mismatches   97;  Indels  321;  Gaps   12;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              || ::  |: | :: :: :||| :|||: |||    |:: || :| ||| ||| ||:|:
Db         27 MCCEHSVAQFSSDQQLNPEENLALYCKPLELYNFIRHRAIENPPYLQRCLLYKIRAKQKK 86

Qy         61 K---------------------------------------------------------- 61
              :
Db         87 RIQITISLPGSNNKELQAQNIFPLYVLFARPTSNVPIEGHSPIYRFSQARLLTSFNDSGN 146

Qy         62 ----------------------------SRST---------------------------- 65
                                          || |
Db        147 NDRAEATFVIPDLETLIATQAYGLTFILVSRGTKKNKGRTGQNLCENDCSEKHVDYSSLR 206

Qy         66 ---------------------------------------------------------- 65

Db        207 KLAGKCFWGKIPITLLNSSLETCADLILGHIVESPISICMSPGYLEPTFLEHDNCLSFCS 266

Qy         66 ---------------------------------------------------------- 65

Db        267 RKADAMVPYQLQVKVSAAEAGAKDILKSPYNSFSYSDVPPSLLLRIVRALQYYANTMKFF 326

Qy         66 ------GMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFK 119
                  | |:|||   || | :|| || :|||| : ||:                   ||
Db        327 LYRLRVGNVLFNYK---NT-QMSEVTEDFTCPFCLVRCGN-----------------FK 364

Qy        120 LFEEYQTVNVSVKLNSFIFEEEGSD-DDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLP 178
              : |:|| |||::| ::   |   :: |:    |   |: :| |               ||
Db        365 ISEDYQAVNVTLKKDNMRTEFVAAEVDNSHRIFYYRSRFKKSRTE------------ILP 412

Qy        179 L---DSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKV 235
              :   |:  : :|:          || |    | |   ::      :: :|    |:
Db        413 VARADAHIMESGS----------------PEET----QAESEDDVQEENENALIDDSKKL 452

Qy        236 ILTSEA---VVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ---------------- 276
                :: :    :   |:||||| |:: |: |||||||| |||: |
Db        453 HGSNHSQSEFLAFGKSRKLSANRADPRNRLLLQKRQFIHSHKAQTPLGDDKVTNLVTADA 512

Qy        277 ---PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVI 333
                 ||  |:|:|| ||||||||||:|| |||:|||||||| |||| ||:|||||||| ::
Db        513 MSVPMTFEEVLSDNDSEDEVDDDIADLEDRRMLDDFVDVTKDEKRIMHMWNSFIRKQSIL 572
```

```
Qy        334 ADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRN 393
              || |: |||||||| : :||   |:| | ||:|:|||||| |: : |:: || ||:: :|
Db        573 ADSHVPWACEAFSRHHGEELLENSALLWGWRMFMIKLWNHSLLSARTMDTCNRILDDIKN 632


RESULT 6
US-10-437-963-114092
; Sequence 114092, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 114092
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1.pep
US-10-437-963-114092

  Query Match             26.0%;  Score 617.5;  DB 6;  Length 201;
  Best Local Similarity   62.6%;  Pred. No. 5.8e-50;
  Matches  114;  Conservative   26;  Mismatches   39;  Indels    3;  Gaps    1;

Qy        223 AIAHSSLDAGAKV---ILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMA 279
              ::||:|:|    :    |::  |:   |||||| ||:: |:  |||||||:|||| ||||
Db         20 SVAHASVDPANSLHGSNLSAPTVLQFGKTRKLSVERADPRNRQLLQKRQFFHSHRAQPMA 79

Qy        280 LEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHIS 339
              |||| |||||||||||:||||||:|||||||| ||||  ||:|||||||||:|||||
Db         80 LEQVFSDRDSEDEVDDDIADFEDRRMLDDFVDVTKDEKLIMHMWNSFVRKQRVLADGHIP 139

Qy        340 WACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTT 399
              |||||||:|: :|| :  :| |||| |::|||||| |:|:  :| ||||||   | |
Db        140 WACEAFSQFHGQELVQNPALLWCWRFFMVKLWNHSLLDARAMNACNTILEGYLNGSSDPK 199

Qy        400 TN 401
              |
Db        200 KN 201


RESULT 7
US-10-424-599-283985
; Sequence 283985, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283985
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1.pep
US-10-424-599-283985

  Query Match             24.2%;  Score 575.5;  DB 6;  Length 242;
  Best Local Similarity   48.2%;  Pred. No. 7e-46;
  Matches  123;  Conservative   38;  Mismatches   63;  Indels   31;  Gaps    6;

Qy        145 DDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGY 204
              | : : |  | || |||    ::    ||   : |:|     ||:  |  :||
Db         13 DPRVQTFFFCGKPLKRRTTADQS----LKNA-VGLESSFPAGGTD--ILEKDDG------ 59

Qy        205 PEATELAGQFEMTSNIPPAIAHSSLDAGAKVILT--SEAVVPATKTRKLSAERSEARSHL 262
                           |   |  |  |  :   ::   :||:   |||||| || : |:
Db         60 ---------------ISATIIRSRPDRDSVQSMSDCDQAVLQFAKTRKLSIERPDPRNST 104

Qy        263 LLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHL 322
              |:||||:|||: ||||:|||:||:|||||||||||||||:||:: |||: ||| |||:
Db        105 FLRKRQFFHSHKAQPMAIEQVLSDKDSEDEVDDDVADFEDRRMLENVVDVSNDEKTFMHM 164

Qy        323 WNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATIN 382
              ||||||| |||||||||||||||: :  |  :  ||  |||:|::||:||||:|: |:|
Db        165 WNSFVRKHRVIADGHISWACEAFSKLHAPEFVQSPSLAGCWRIFMVKLYNHGLLDARTMN 224

Qy        383 NCNTILENC-RNSSD 396
              :|| |||   | :||
Db        225 DCNIILEQYQRQNSD 239


RESULT 8
US-10-198-723-6
; Sequence 6, Application US/10198723
; GENERAL INFORMATION:
;  APPLICANT: Dang, Van-Dinh
;  APPLICANT:  Okamuro, Jack
;  TITLE OF INVENTION: CHIMERIC POLYPEPTIDES MODULATING
;  TITLE OF INVENTION:  DEVELOPMENT
;  FILE REFERENCE: 11696-043001
;  CURRENT APPLICATION NUMBER: US/10/198,723
;  CURRENT FILING DATE:  2002-07-05
;  PRIOR APPLICATION NUMBER: US 60/303,654
;  PRIOR FILING DATE: 2001-07-06
;  NUMBER OF SEQ ID NOS: 100
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 692
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-198-723-6

  Query Match             21.5%;  Score 510.5;  DB 6;  Length 692;
  Best Local Similarity   22.3%;  Pred. No. 4.1e-39;
  Matches  154;  Conservative   53;  Mismatches  121;  Indels  363;  Gaps    8;

Qy         81 KTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVSVKLNSFIF 138
              | || |: ||||| : ||  :||| || |||| |:|||   |:     |:|||| ::  |
Db          4 KAEVVENFSCPFCLIPCGGHEGLQLHLKSSHDAFKFEFYRAEKDHGPEVDVSVKSDTIKF 63

Qy        139 E--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195
                 ::   : :  | : ||| | :||||   || ::| |  : ||   |  |||| | :
Db         64 GVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPRGTENDSTHV 123

Qy        196 NDGN---------------------------------------------------------- 199
              || |
Db        124 NDDNVSSPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAHSSAEKNEST 183

Qy        200 ---------------------------------------------------------- 199

Db        184 HVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDEDVSFPPRTRS 243

Qy        200 ------------------------RG------------------------------- 201
                                       ||
Db        244 SKETSDILTTTQPAIVEPSEPKVRRGSRRKQLYAKRYKARETQPAIAESSEPKVLHVNDE 303

Qy        202 --LGYPEATELAGQFEMTSNIPPAIAHSS---------------------------- 228
                  |||  |    :: :   |||| ||
```

```
Db        304 NVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSKKNKSTRKNV 363

Qy        229 ------------------------------------------------------------ 228

Db        364 DNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSSKKNKSTRKN 423

Qy        229 -------------------LDAGAKVILTSEAVVP------------------------- 244
                                  |      |  ||  ||
Db        424 DDNIPSPPKTRSSKKTSNILTRTQPAIAESEPKVPHVNDDKVSSTPRAHSSKKNKSTHKK 483

Qy        245 -----------------------------------ATKTRKLSAERSEARSHLLLQKR 267
                                                  :: ::| ||| ||:    |: |
Db        484 DDNASLPPKTRSSKKTSDILATTQPAKAEPSEPKVTRVSRRKELHAERCEAKRLERLKGR 543

Qy        268 QFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFV 327
              |||||  :|||  |||||: |||:| ||   |  :|  |:  | |:|:||::|:|||| ||
Db        544 QFYHSQTMQPMTFEQVMSNEDSENETDDYALDISERLRLERLVGVSKEEKRYMYLWNIFV 603

Qy        328 RKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTI 387
              ||||||||||: |||| |:: :::|:   ||  | ||:| |||||:||: : | : | ||
Db        604 RKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDWWWRMFRIKLWNNGLICAKTFHKCTTI 663

Qy        388 LENCRNSSDTTTTNNNNSVDRPSDSNTNNNN 418
              |          :|:: :    | |  | ||
Db        664 L----------LSNSDEAGQFTSGSAANANN 684


RESULT 9
US-10-424-599-254263
; Sequence 254263, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254263
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(166)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1.pep
US-10-424-599-254263

  Query Match             20.6%;  Score 491;  DB 6;  Length 166;
  Best Local Similarity   60.2%;  Pred. No. 4e-38;
  Matches  100;  Conservative   21;  Mismatches   39;  Indels    6;  Gaps    3;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||||     : || |: ||:|||||||| |||| : |:| ||||| ||| ||| |:|||
Db          1 MCRQNSPVHHAGEEEIAADESLLIYCKPVELYNILYRRALQNPSFLRRCLLYKIRARRKR 60

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
              : |: |:|:|||:|  | |:|||| || ||||| | ||||||||:||| |||||| | | :
Db         61 RLRA-GIVIFNYRDHYNILRKTEVTEDFSCPFCLMQCGSFKGLRFHLCSSHDLFNFGFWV 119

Qy        121 FEEYQTVNVSVKLNSFIFEEEGSDD---DKFEPFSLCSKPRKRRQR 163
               |:|| ||||||:|  ||| |   |    : : |  ||:|| ||::
Db        120 TEDYQAVNVSVKIN--IFEXENVADGVIPQSQTFFFCSRPRNRRRK 163


RESULT 10
US-10-424-599-183487
; Sequence 183487, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 183487
;   LENGTH: 102
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(102)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1.pep
US-10-424-599-183487

  Query Match             11.3%;  Score 269;  DB 6;  Length 102;
  Best Local Similarity   52.1%;  Pred. No. 1.9e-17;
  Matches   50;  Conservative   14;  Mismatches   32;  Indels    0;  Gaps    0;

Qy        289 SEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRF 348
              |: ||| |  | ||   |     |:|||| |||  :| :||: |   |:: |||||| :|
Db          7 SQQEVDHDPDDVEDPATLHHLAHVSKDEKXFMHXCHSLMRKKSVCPHGYVXWACEAFXQF 66

Qy        349 YEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNC 384
              | |||    :||| ||||:|||||:||: : |:| |
Db         67 YGKELILSLALFWRWRLFMIKLWNYGLLHARTMNXC 102


RESULT 11
US-10-424-599-216298
; Sequence 216298, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216298
;   LENGTH: 116
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_37346C.1.pep
US-10-424-599-216298

  Query Match             10.8%;  Score 256.5;  DB 6;  Length 116;
  Best Local Similarity   41.2%;  Pred. No. 3.4e-16;
  Matches   47;  Conservative   26;  Mismatches   38;  Indels    3;  Gaps    2;

Qy        293 VDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQ--RVIADGHISWACEAFSRFYE 350
              :::   : |||: |   :|     |:|: |||:|::|   ||:|||||:||||||::::
Db          1 MNEGAQEIEDRRRL-AILDAPDHVKEFLSLWNAFIKKHRGRVLADGHINWACEAFTKYHS 59

Qy        351 KELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNN 404
              |  : :|| | ||:|:||| | ||: ::|   || ||| ||  :      | :
Db         60 AEFAQSNSLAWNWRMFMIKLINQGLLKTSTFVACNNILEQCRKQNQDPQRQNQD 113


RESULT 12
US-10-437-963-199908
```

```
; Sequence 199908, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199908
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_9542C.1.pep
US-10-437-963-199908

  Query Match             10.7%;  Score 255.5;  DB 6;  Length 151;
  Best Local Similarity   42.4%;  Pred. No. 6.1e-16;
  Matches   67;  Conservative   15;  Mismatches   39;  Indels   37;  Gaps    5;

Qy          4 QNCRAK---SSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |   ||    ||  | | ::       |  | :|  |||
Db         25 QEAGAKDMSESPYSVYSYND-----VPPSSLTHIIRLRS---------------------- 58

Qy         61 KSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKL 120
                | |:|||  |||:||||| || ||||| : ||||||  |||:||||| :|| :
Db         59 -----GNVLFNYKYYNNTMQKTEVTEDFSCPFCLVPCGSFKGLGCHLNASHDLFHYEFWI 113

Qy        121 FEEYQTVNVSVKLNSFIFE--EEGSDDDKFEPFSLCSK 156
              || | ||||:| :|:  |   ||  | : : ||  |:
Db        114 SEECQAVNVSLKTDSWRTELLAEGV-DPRHQTFSYRSR 150


RESULT 13
US-10-425-114-70148
; Sequence 70148, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70148
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73067D07_FLI.pep
US-10-425-114-70148

  Query Match             10.5%;  Score 249;  DB 6;  Length 394;
  Best Local Similarity   20.8%;  Pred. No. 9.5e-15;
  Matches   70;  Conservative   17;  Mismatches   29;  Indels  220;  Gaps    2;

Qy          3 RQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR-- 60
              ||  |   ||:| :: :::| :||||| ||||   |:: || |: ||| | | |:||:
Db         55 RQQLRTALSPDEKLTAEKDLALYCKPVELYNIIQRRAMKNPLFIQRCLLYNIHARRKKRI 114

Qy         61 ------------------------------------------------------------ 60

Db        115 QITISLSGSTNTELQTHYVFPLYVLLARPTSNLSLEGHSPIYRFSRVCLLTSFSEHGNKD 174

Qy         61 -----------KSRST---------------------------------------------- 65
                         || ||
Db        175 NSEATFIIPDVKSLSTSRACNHDIIFISCGQVGQSNGEDNCSGNHVEDSSLQMLEGKCSW 234

Qy         66 ------------------------------------------------------------ 65

Db        235 GKIPTNLLASSLESCVNLSLGHIVELASKVTMRPSFLEPKFLEQDSCLTFCSHKVDAVGS 294

Qy         66 -----------------------------------------GMVVFNYKDCNNTLQKT 82
                                                        | |:|||| .|||:|:|
Db        295 YKLQLCMSAQEAGARDMSLSPYSSYSYNDVPPSSLSDIIRLRSGNVLFNYKYYNNTMQET 354

Qy         83 EVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEF 118
              || || ||||| : ||||||| |||||||| :||
Db        355 EVTEDFSCPFCYVRCGSFKGLGCHLNSSHDLFHYEF 390


RESULT 14
US-10-437-963-184120
; Sequence 184120, Application US/10437963
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184120
;   LENGTH: 263
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(263)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_81143C.1.pep
US-10-437-963-184120

  Query Match             8.0%;  Score 190;  DB 6;  Length 263;
  Best Local Similarity   53.8%;  Pred. No. 2e-09;
  Matches   35;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKR 60
              |||   ||: ||:| :: :|:  :||||| ||||   ||: ||:|| ||| ||| |:|:
Db          3 MCRHQPRARLSPDEQLAAEESFALYCKPVELYNIIQRRSIKNPAFLQRCLLYKIHARRKK 62

Qy         61 KSRST 65
              :|  |
Db         63 RSLIT 67


RESULT 15
US-10-424-599-250979
; Sequence 250979, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250979
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(195)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_68663C.1.pep
US-10-424-599-250979

  Query Match             6.7%;  Score 159;  DB 6;  Length 195;
  Best Local Similarity   56.9%;  Pred. No. 1.1e-06;
  Matches   29;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         11 SPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRK 61
              | || :: :|:| |||||| ||||   |:: ||||| |||:|:| || |::
Db         22 SAEEELAAEESLSIYCKPVELYNILQRRAMRNPSFLQRCLHYRIKAKHKKR 72


Search completed: June 13, 2003, 15:37:12
Job time : 56 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 22:10:23 ; Search time 3586 Seconds
(without alignments)
3611.475 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144851_16477/app_query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_2496_@runat_13062003_144851_16477 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2378	100.0	1722	6	AX032817	AX032817 Sequence
	2	2378	100.0	1722	6	AX032888	AX032888 Sequence
	3	2370	99.7	1700	8	AF284500	AF284500 Arabidops
	4	2363	99.4	1722	6	AX032934	AX032934 Sequence
	5	2343.5	98.5	1722	8	AF284501	AF284501 Arabidops
	6	2285	96.1	1354	8	AY063047	AY063047 Arabidops
	7	2285	96.1	1715	6	AX032818	AX032818 Sequence
	8	2285	96.1	1715	6	AX032891	AX032891 Sequence
	9	2285	96.1	1721	8	AY034902	AY034902 Arabidops
	10	2258.5	95.0	1737	6	AX032894	AX032894 Sequence
	11	1813.5	76.3	6338	6	AX032890	AX032890 Sequence
c	12	1798.5	75.6	91660	8	ATRPP5LE2	AF180942 Arabidops
	13	1659	69.8	5895	6	AX032893	AX032893 Sequence
	14	1659	69.8	199548	8	ATCHRIV45	AL161545 Arabidops
	15	1659	69.8	201471	8	ATFCA7	Z97342 Arabidopsis
	16	923	38.8	2279	8	AB053171	AB053171 Arabidops
	17	589	24.8	2384	8	AF096096	AF096096 Arabidops
	18	451.5	19.0	558	6	AX032900	AX032900 Sequence
	19	400.5	16.8	8508	8	AB053262	AB053262 Arabidops
	20	400.5	16.8	87180	8	AB023044	AB023044 Arabidops
	21	368.5	15.5	325	6	AX032898	AX032898 Sequence
	22	311	13.1	463	6	AX032896	AX032896 Sequence
c	23	306.5	12.9	135495	2	OSJN01007	AL607099 Oryza sat
c	24	265	11.1	107436	8	AC006068	AC006068 Arabidops
	25	264.5	11.1	6081	8	AF096095	AF096095 Arabidops
	26	250	10.5	2040	8	AB053265	AB053265 Arabidops
c	27	250	10.5	199548	8	ATCHRIV45	AL161545 Arabidops
c	28	250	10.5	201471	8	ATFCA7	Z97342 Arabidopsis
	29	242	10.2	127130	8	OSJN00040	AL606603 Oryza sat
c	30	240	10.1	113355	8	OSJN00100	AL606654 Oryza sat
	31	222	9.3	3346	9	AK074333	AK074333 Homo sapi
	32	221	9.3	3942	9	BC015704	BC015704 Homo sapi
	33	221	9.3	4441	6	AX032902	AX032902 Sequence
	34	221	9.3	4441	6	AX337789	AX337789 Sequence
	35	221	9.3	4441	9	D63881	D63881 Human mRNA
	36	219	9.2	3648	3	AF149047	AF149047 Drosophil
	37	219	9.2	4055	3	AY069809	AY069809 Drosophil
c	38	208	8.7	83538	2	AC018338	AC018338 Drosophil
	39	208	8.7	167816	3	AC009381	AC009381 Drosophil
c	40	208	8.7	249805	3	AE003515	AE003515 Drosophil
c	41	178.5	7.5	57873	2	AC010021	AC010021 Drosophil
	42	175	7.4	56152	2	AC116963	AC116963 Dictyoste
	43	168	7.1	5769	3	AF104350	AF104350 Dictyoste
c	44	161	6.8	171187	2	AC116960	AC116960 Dictyoste
c	45	155.5	6.5	12464	2	AC115604	AC115604 Dictyoste

ALIGNMENTS

RESULT 1

```
AX032817
LOCUS       AX032817               1722 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 57 from Patent WO0046358.
ACCESSION   AX032817
VERSION     AX032817.1  GI:10279793
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Dean,C., West,J. and Johanson,U.
  TITLE     Plant gene
  JOURNAL   Patent: WO 0046358-A 57 10-AUG-2000;
            DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
            JOHANSON URBAN (SE)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Landsberg erecta VRN2 cDNA"
BASE COUNT      538 a    339 c    348 g    497 t
ORIGIN

Alignment Scores:
Pred. No.:              1.46e-209       Length:         1722
Score:                  2378.00         Matches:        445
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-890-220-2 (1-445) x AX032817 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 2
AX032888
LOCUS       AX032888               1722 bp    DNA     linear   PAT 22-SEP-2000
DEFINITION  Sequence 1 from Patent WO0044918.
ACCESSION   AX032888
VERSION     AX032888.1  GI:10279825
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 1 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1722
```

```
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT       538 a    339 c    348 g    497 t
ORIGIN

Alignment Scores:
Pred. No.:                1.46e-209      Length:          1722
Score:                    2378.00        Matches:         445
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-890-220-2 (1-445) x AX032888 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              ||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 3
AF284500
LOCUS       AF284500               1700 bp    mRNA    linear   PLN 21-NOV-2001
DEFINITION  Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
            cds.
ACCESSION   AF284500
VERSION     AF284500.1  GI:16945787
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1700)
  AUTHORS   Gendall,A.R., Levy,Y.Y., Wilson,A. and Dean,C.
  TITLE     The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of
            Vernalization in Arabidopsis
  JOURNAL   Cell 107 (4), 525-535 (2001)
  MEDLINE   21575875
   PUBMED   11719192
REFERENCE   2  (bases 1 to 1700)
  AUTHORS   Gendall,A.R., Levy,Y.Y. and Dean,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre,
            Colney Lane, Norwich NR4 7UH, UK
FEATURES             Location/Qualifiers
     source          1. .1700
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="adjacent to RPP5"
                     /note="ecotype: Lansdberg erecta"
     gene            1. .1700
                     /gene="VRN2"
     CDS             227. .1564
                     /gene="VRN2"
                     /codon_start=1
                     /product="vernalization 2 protein"
                     /protein_id="AAL32135.1"
```

```
                     /db_xref="GI:16945788"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKG
                     LQFHLNSSHDLFEFEFKLSEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
                     RQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTS
                     NIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
                     PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
                     DGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
                     NSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK"
     variation       853
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="a"
     variation       1195
                     /gene="VRN2"
                     /note="vrn2-1 mutation results in TGG Trp to TGA stop"
                     /replace="a"
     variation       1399
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="g"
     variation       1487
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="a"
     variation       1488
                     /gene="VRN2"
                     /note="compared to the sequence in GenBank Accesssion
                     Number AF180942"
                     /replace="g"
BASE COUNT      533 a    334 c    347 g    486 t
ORIGIN

Alignment Scores:
Pred. No.:              7.83e-209       Length:         1700
Score:                  2370.00         Matches:        444
Percent Similarity:     99.78%          Conservative:   0
Best Local Similarity:  99.78%          Mismatches:     1
Query Match:            99.66%          Indels:         0
DB:                     8               Gaps:           0

US-09-890-220-2 (1-445) x AF284500 (1-1700)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 286

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 346

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 406

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 466

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 526

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 586

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 TCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 646

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 706

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 766

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 826

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 886

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 946

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1006

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1066

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1126

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1186

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATTTCGTGG 1246

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1306

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1366

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1547 GACAAAGTAATTAAA 1561

RESULT 4
AX032934
LOCUS       AX032934               1722 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 47 from Patent WO0044918.
ACCESSION   AX032934
VERSION     AX032934.1  GI:10279842
KEYWORDS    .
SOURCE      thale cress.
```

```
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 47 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      539 a    339 c    347 g    497 t
ORIGIN

Alignment Scores:
Pred. No.:                3.52e-208      Length:          1722
Score:                    2363.00        Matches:         444
Percent Similarity:       99.78%         Conservative:    0
Best Local Similarity:    99.78%         Mismatches:      1
Query Match:              99.37%         Indels:          0
DB:                       6              Gaps:            0

US-09-890-220-2 (1-445) x AX032934 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGAAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 5
AF284501
LOCUS       AF284501               1722 bp    mRNA    linear   PLN 21-NOV-2001
DEFINITION  Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
            cds, alternatively spliced.
ACCESSION   AF284501
VERSION     AF284501.1  GI:16945789
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Gendall,A.R., Levy,Y.Y., Wilson,A. and Dean,C.
  TITLE     The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of
            Vernalization in Arabidopsis
  JOURNAL   Cell 107 (4), 525-535 (2001)
  MEDLINE   21575875
   PUBMED   11719192
REFERENCE   2  (bases 1 to 1722)
  AUTHORS   Gendall,A.R., Levy,Y.Y. and Dean,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre,
            Colney Lane, Norwich NR4 7UH, UK
FEATURES             Location/Qualifiers
```

```
     source          1. .1722
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="IV"
                     /map="adjacent to RPP5"
                     /note="ecotype: Lansdberg erecta"
     gene            1. .1722
                     /gene="VRN2"
     CDS             227. .550
                     /gene="VRN2"
                     /note="alternatively spliced"
                     /codon_start=1
                     /product="vernalization 2 protein"
                     /protein_id="AAL32136.1"
                     /db_xref="GI:16945790"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKV
                     GNYYN"
BASE COUNT      540 a    338 c    353 g    491 t
ORIGIN

Alignment Scores:
Pred. No.:                2.21e-206      Length:          1722
Score:                    2343.50        Matches:         444
Percent Similarity:       98.01%         Conservative:    0
Best Local Similarity:    98.01%         Mismatches:      1
Query Match:              98.55%         Indels:          8
DB:                       8              Gaps:            1

US-09-890-220-2 (1-445) x AF284501 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 286

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 346

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 406

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 466

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 526

Qy        101 Lys----------------------GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
              |||                      ||||||||||||||||||||||||||||||||||||
Db        527 AAGGTGGGCAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTT 586

Qy        113 uPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db        587 ATTTGAATTTGAGTTCAAGCTTTCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACT 646

Qy        133 uAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSerLe 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 TAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCT 706

Qy        153 uCysSerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLy 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 CTGCTCGAAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAA 766

Qy        173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 AGTATGCTTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCAC 826

Qy        193 rLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGl 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 CCTACTTAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACA 886

Qy        213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ATTTGAGATGACCAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGC 946

Qy        233 aLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAl 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TAAAGTTATATTGACAAGCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGC 1006

Qy        253 aGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHi 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCA 1066

Qy        273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVa 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGT 1126

Qy        293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 CGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAA 1186

Qy        313 nLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIl 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 TAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTAT 1246

Qy        333 eAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLe 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 AGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTT 1306

Qy        353 uHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHi 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCA 1366

Qy        373 sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 TGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAA 1426

Qy        393 nSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAs 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAA 1486

Qy        413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACCAACAACAATAACATTGTGGATCATCCCAATGACATAAACAACAAGAACAATGTTGA 1546

Qy        433 pAsnLysAspAsnAsnSerArgAspLysValIleLys 445
              |||||||||||||||||||||||||||||||||||||
Db       1547 CAACAAGGACAATAACAGCAGAGACAAAGTAATTAAA 1583

RESULT 6
AY063047
LOCUS       AY063047               1354 bp    mRNA    linear   PLN 18-APR-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
            cds.
ACCESSION   AY063047
VERSION     AY063047.1  GI:17104664
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1354)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Etgu,P.,
            Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H.,
            Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
            Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
            Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1354)
  AUTHORS   Yamada,K., Banh,J., Banno,F., Chang,E., Dale,J.M., Goldsmith,A.D.,
```

Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Banh,J., Banno,F., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PGEC) contributed equally to this work as PIs.

FEATURES Location/Qualifiers

source 1. .1354
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U09871"
/note="This clone is in pUNI 51.
ecotype: Columbia"

gene 1. .1354
/gene="At4g16845"

CDS 1. .1323
/gene="At4g16845"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AAL34221.1"
/db_xref="GI:17104665"
/translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKG
LQFHLNSSHDLFEFEFKLLEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
RQRGGRNNTRRLKVCFLPLDSPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTS
NIPPAIAHSSLDAGAKVILTTEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
DGHISWACEVFSRFYEKELHCYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
NTSVTNNNNNSVDHPSDSNTNNNNIVDHPNDIKNKNNVDNKDNNSRDK"

3'UTR 1324. .1354
/gene="At4g16845"

BASE COUNT 431 a 264 c 287 g 372 t

ORIGIN

Alignment Scores:

Pred. No.:	3.94e-201	Length:	1354
Score:	2285.00	Matches:	428
Percent Similarity:	97.51%	Conservative:	3
Best Local Similarity:	96.83%	Mismatches:	9
Query Match:	96.09%	Indels:	2
DB:	8	Gaps:	1

US-09-890-220-2 (1-445) x AY063047 (1-1354)

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 60

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 120

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 180

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 240

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 300

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 360

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 420

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 480

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 540

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              |||||||||||  ||||||||||||||||||||  ||||||||||||||||||||||||
Db        541 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 600

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 660

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db        661 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 720

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 780

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 840

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 900

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 960

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1020

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||
Db       1021 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1080

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1140

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||:::|||         |||
```

```
Db       1141 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1194

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       1195 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1254

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       1255 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1314

Qy        441 AspLys 442
              ||||||
Db       1315 GACAAG 1320

RESULT 7
AX032818
LOCUS       AX032818                1715 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 58 from Patent WO0046358.
ACCESSION   AX032818
VERSION     AX032818.1  GI:10279794
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1715)
  AUTHORS   Dean,C., West,J. and Johanson,U.
  TITLE     Plant gene
  JOURNAL   Patent: WO 0046358-A 58 10-AUG-2000;
            DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
            JOHANSON URBAN (SE)
FEATURES             Location/Qualifiers
     source          1. .1715
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Columbia VRN2 cDNA"
BASE COUNT      536 a    326 c    353 g    500 t
ORIGIN

Alignment Scores:
Pred. No.:                5.42e-201      Length:          1715
Score:                    2285.00        Matches:         428
Percent Similarity:       97.51%         Conservative:    3
Best Local Similarity:    96.83%         Mismatches:      9
Query Match:              96.09%         Indels:          2
DB:                       6              Gaps:            1

US-09-890-220-2 (1-445) x AX032818 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||   ||||||||||||||||||||   ||||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||   ||||||||||||||||||||||||||||   ||||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||:::|||         |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLys 442
              ||||||
Db       1547 GACAAG 1552

RESULT 8
AX032891
LOCUS       AX032891                1715 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 4 from Patent WO0044918.
ACCESSION   AX032891
VERSION     AX032891.1  GI:10279827
```

```
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1715)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 4 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1715
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      536 a    326 c    353 g    500 t
ORIGIN

Alignment Scores:
Pred. No.:              5.42e-201       Length:         1715
Score:                  2285.00         Matches:        428
Percent Similarity:     97.51%          Conservative:   3
Best Local Similarity:  96.83%          Mismatches:     9
Query Match:            96.09%          Indels:         2
DB:                     6               Gaps:           1

US-09-890-220-2 (1-445) x AX032891 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||   ||||||||||||||||||||   ||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||   ||||||||||||||||||||||||||||||   ||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||:::|||         |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLys 442
              ||||||
Db       1547 GACAAG 1552

RESULT 9
AY034902
LOCUS       AY034902               1721 bp    mRNA    linear   PLN 23-APR-2002
DEFINITION  Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
            cds.
ACCESSION   AY034902
VERSION     AY034902.1  GI:14334421
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1721)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1721)
```

```
AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
          Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G.,
          Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
          Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
          Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M.,
          Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
          Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
          Theologis,A.
TITLE     Direct Submission
JOURNAL   Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan
          Street, Albany, CA 94710, USA
COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
          collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
          Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
          Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
          Hayashizaki,Y. and Shinozaki,K.

          The Salk, Stanford, PGEC (SSP) Consortium members carried out the
          sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X.,
          Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D.,
          Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L.,
          Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E.,
          Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J.,
          Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

          Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
          this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
          contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1721
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL05-08-K19 (R09871)"
                     /note="This clone is in a modified pBluescript vector
                     (Lambda ZAP) as a XhoI/SstI insert.
                     ecotype: Columbia"
     gene            1. .1721
                     /gene="At4g16845"
     5'UTR           1. .231
                     /gene="At4g16845"
     CDS             232. .1554
                     /gene="At4g16845"
                     /codon_start=1
                     /evidence=experimental
                     /product="unknown protein"
                     /protein_id="AAK59409.1"
                     /db_xref="GI:14334422"
                     /translation="MCRQNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPS
                     FLPRCLNYKIGAKRKRKSRSTGMVVFNYKDCNNTLQRTEVREDCSCPFCSMLCGSFKG
                     LQFHLNSSHDLFEFEFKLLEEYQTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKR
                     RQRGGRNNTRRLKVCFLPLDSPSLANGTENGIALLNDGNRGLGYPEATELAGQFEMTS
                     NIPPAIAHSSLDAGAKVILTTEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQ
                     PMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIA
                     DGHISWACEVFSRFYEKELHCYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCR
                     NTSVTNNNNNSVDHPSDSNTNNNNIVDHPNDIKNKNNVDNKDNNSRDK"
     3'UTR           1555. .1721
                     /gene="At4g16845"
BASE COUNT      548 a    324 c    353 g    496 t
ORIGIN

Alignment Scores:
Pred. No.:                5.45e-201     Length:          1721
Score:                    2285.00       Matches:         428
Percent Similarity:       97.51%        Conservative:    3
Best Local Similarity:    96.83%        Mismatches:      9
Query Match:              96.09%        Indels:          2
DB:                       8             Gaps:            1

US-09-890-220-2 (1-445) x AY034902 (1-1721)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        232 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 291

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        292 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 351

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        352 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 411

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        412 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 471

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        472 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 531

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        532 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 591

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        592 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 651

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        652 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 711

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        712 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 771

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||

Db        772 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 831

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        832 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 891

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||

Db        892 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 951

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        952 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1011

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1012 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1071

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1072 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1131

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1132 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1191

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1192 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1251

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||

Db       1252 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1311

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1312 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1371
```

```
Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||:::|||          |||
Db       1372 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1425

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1426 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1485

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1486 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1545

Qy        441 AspLys 442
              ||||||
Db       1546 GACAAG 1551

RESULT 10
AX032894
LOCUS       AX032894               1737 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 7 from Patent WO0044918.
ACCESSION   AX032894
VERSION     AX032894.1  GI:10279829
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1737)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 7 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .1737
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT      543 a    330 c    359 g    505 t
ORIGIN

Alignment Scores:
Pred. No.:              1.53e-198       Length:         1737
Score:                  2258.50         Matches:        428
Percent Similarity:     95.78%          Conservative:   3
Best Local Similarity:  95.11%          Mismatches:     9
Query Match:            94.97%          Indels:         10
DB:                     6               Gaps:           2

US-09-890-220-2 (1-445) x AX032894 (1-1737)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 Lys----------------------GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
              |||                     ||||||||||||||||||||||||||||||||||||
Db        533 AAGGTGGGCAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTT 592

Qy        113 uPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        593 ATTTGAATTTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACT 652

Qy        133 uAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSerLe 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCT 712

Qy        153 uCysSerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLy 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAA 772

Qy        173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh 193
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        773 AGTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGC 832

Qy        193 rLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGl 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 CCTGCTGAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACA 892

Qy        213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATTTGAGATGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGC 952

Qy        233 aLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAl 253
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        953 TAAAGTTATATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGC 1012

Qy        253 aGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHi 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCA 1072

Qy        273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVa 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 CAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGT 1132

Qy        293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAA 1192

Qy        313 nLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIl 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTAT 1252

Qy        333 eAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLe 353
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db       1253 AGCAGATGGTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTT 1312

Qy        353 uHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHi 373
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 GCACTGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCA 1372

Qy        373 sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAA 1432

Qy        393 nSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAs 413
              |:::|||         |||   ||||||||||||||||||   ||||||||||||||
Db       1433 TACCTCA------GTCACTAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAA 1486

Qy        413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs 433
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db       1487 CACCAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGA 1546

Qy        433 pAsnLysAspAsnAsnSerArgAspLys 442
              ||||||||||||||||||||||||||||
Db       1547 CAACAAGGACAATAACAGCAGAGACAAG 1574

RESULT 11
```

```
AX032890
LOCUS       AX032890               6338 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 3 from Patent WO0044918.
ACCESSION   AX032890
VERSION     AX032890.1  GI:10279826
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 6338)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 3 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .6338
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT     1932 a   1076 c   1134 g   2193 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:              9.1e-157        Length:         6338
Score:                  1813.50         Matches:        440
Percent Similarity:     43.58%          Conservative:   1
Best Local Similarity:  43.48%          Mismatches:     4
Query Match:            76.26%          Indels:         569
DB:                     6               Gaps:           13

US-09-890-220-2 (1-445) x AX032890 (1-6338)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 2401

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2461

Qy         41 GlyAsn---------------------------------------------------------- 42
              ||||||
Db       2462 GGCAACGTATGATTTGCCTTCCTCTCTCATCATATTAGCTCAGTAATCTTTCATCTCCTG 2521

Qy         42 ---------------------------------------------------------------- 42

Db       2522 TGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGATTCATGGCGA 2581

Qy         43 ----------------------------------------------------ProSerPhe 45
                                                                  ||||||||
Db       2582 GTGTGTGCTTCTTTTGTCTCCTAAATTTGAACTTGTTGTTTGTTGTTGCAGCCATCGTTT 2641

Qy         46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer---------- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 CTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTA-TGTTTTCTT 2700

Qy         62 ---------------------------------------------------------------- 62

Db       2701 CTTGAATGTAGCTGCTACAGTGATATGTTATTTATCTTACTTCTAATATGGAAGCTGATG 2760

Qy         62 ---------------------------------------------------------------- 62

Db       2761 ACCTATTTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTGTTCATG 2820

Qy         63 ----------------------------------------ArgSerThrGlyMetVal 68
                                                      ||||||||||||||||||
Db       2821 CTATAAACTTACATTTTATAAAATTGTGTTTTGGTTAGGTCAAGATCTACTGGGATGGTA 2880

Qy         69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------------- 83
              |||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGGTTAGTCTTTTTCT 2940

Qy         83 ---------------------------------------------------------------- 83

Db       2941 GTTCGTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTATTTACTA 3000

Qy         84 ----------------------ValArgGluAspCysSerCysProPheCysSerMetLe 96
                                    ||||||||||||||||||||||||||||||||||||
Db       3001 TTTTTTTCTGTATTGTCATGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCT 3060

Qy         96 uCysGlySerPheLys------------------------------------------- 101
              ||||||||||||||||
Db       3061 ATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCTTTCATA 3120

Qy        101 ---------------------------------------------------------------- 101

Db       3121 TCTAACACTGTGAAATGCTACTGCTGTTTCATGCTGTATACTTTCACTGTTTGGTTACAT 3180

Qy        101 ---------------------------------------------------------------- 101

Db       3181 ATTTTTGTGTTTGTTGTTTGTCTTCTCACTCTTTTCGAACTGCTGAGTGTGTGCTTATCT 3240

Qy        102 ----------------------------------------------------------GlyLeuG 104
                                                                         |||||||
Db       3241 GAGAAAACATGTCCCAGATGGAGCTTACAACCAATTGTCTTGTGTCTATGCAGGGGCTGC 3300

Qy        104 lnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu---------- 120
              ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       3301 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATGTGGTTTTA 3360

Qy        120 ---------------------------------------------------------------- 120

Db       3361 TGGAAATTCTTGATTTGCTATGCCTTTATTAATGAGGTTATAGTTAAAAAAGGGTCTTTC 3420

Qy        121 -----------Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
                          ||| ||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CTATTGTAGCTTTCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTC 3480

Qy        137 IlePheGlu--------------------------------------------------- 139
              ||||||||
Db       3481 ATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCGATAGCTGGTGAAAAT 3540

Qy        140 --------------------------------GluGluGlySerAspAspAspLysPheGl 149
                                              ||||||||||||||||||||||||||||
Db       3541 TTTGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGACGATAAATTTGA 3600

Qy        149 uProPheSerLeu---------------------------------------------- 153
              ||||||||||||
Db       3601 GCCCTTCTCTCTCTGGTAACCCTCAGAACCCCTTCGATTAAATACCTTAATAGCAGTAAC 3660

Qy        154 ----------------------------------------------------CysSerLy 156
                                                               |||||
Db       3661 TCCTTGCTTCTCTTGTCAGTACATCTCTGTAAATCCAACCATAATGTTTTGCAGCTCGAA 3720

Qy        156 sProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPh 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTT 3780

Qy        176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAs 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAA 3840

Qy        196 nAsp------------------------------------------------------- 197
              ||||
Db       3841 TGATGGTAAAATCATATCTTCTTCTGTGCGTTCCTTGTGGCTTAGAACTTCATATTACAG 3900

Qy        197 ---------------------------------------------------------------- 197

Db       3901 AAGAAGATACAATGGCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTCTTCTTGCGA 3960

Qy        198 ----------------------------------------------GlyAsnArgGlyLe 202
                                                             ||||||||||||||
Db       3961 GGGTATTGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAACCGTGGTTT 4020
```

```
Qy       202 uGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPr 222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4021 AGGATATCCCGAGGCAACAGARCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACC 4080

Qy       222 oAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVa 242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4081 AGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGT 4140

Qy       242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg-------- 259
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4141 GGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGGTTTGT 4200

Qy       259 ---------------------------------------------------------- 259

Db      4201 TCATCATGACACCCCGTCATCATAATTACCATTCCTGTTGTTACAAATGTTCTTCCTATT 4260

Qy       260 ----------------------------------------------------SerHisLeuLe 263
                                                                 ||||||||||||
Db      4261 ATGGATAAGTGTTTATAGTACTGCCATATTAACCGAGAAAATTTCTTCCAGCCACCTACT 4320

Qy       263 uLeuGlnLysArgGlnPheTyrHisSerHisArgVal----------------------- 275
            ||||||||||||||||||||||||||||||||||||||
Db      4321 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGTTCCTTCACC 4380

Qy       275 ---------------------------------------------------------- 275

Db      4381 TACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGCATACTTGGT 4440

Qy       275 ---------------------------------------------------------- 275

Db      4441 TTGTTCTCATCCAAATTTGTATTTTGATCTGTATGTATCAACGCAAAATAGTTATGTCCA 4500

Qy       276 ----------------------------------------------------Gln-ProMetA 279
                                                                 ||| |||||||
Db      4501 TGTTGTCTCCGTTTTATTGCCACTAACCAAAAAATGCATGTTTCTGTGACAAGCCAATGG 4560

Qy       279 laLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaA 299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4561 CGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAG 4620

Qy       299 spPheGluAspArg----------------------------------------------- 303
            |||||||||||||
Db      4621 ATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAATAGACAACAGA 4680

Qy       304 ----------------------------------------------------GlnMetLeu 306
                                                                 ||||||||
Db      4681 AAATGGTATATGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGCTGCAGATGCTT 4740

Qy       307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4741 GATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTT 4800

Qy       327 ValArgLysGlnArgVal------------------------------------------ 332
            ||||||||||||||||
Db      4801 GTAAGAAAACAAAGGTA-ACTACTTCTCTTACACATGAACAGACACAAAAAGACCTTATG 4859

Qy       332 ---------------------------------------------------------- 332

Db      4860 TCTTACATTCCATACCTGTCTAAATGATTTTGCTTATGGAACTTTGAGCTCAATTATGAT 4919

Qy       333 ------------------IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346
                              ||||||||||||||||||||||||||||||||||||||||||
Db      4920 TGTTGATGTTTCAGGGTTATAGCAGATGGTCATATTTCTTGGGCATGTGAAGCATTTTCA 4979

Qy       347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp--------------- 361
            |||||||||||||||||||||||||||||||||||||||||||
Db      4980 AGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTAATATAAGTACACC 5039

Qy       361 ---------------------------------------------------------- 361

Db      5040 AAACATATACAGACACATAACTACACTATCAATCTTGTTTCGTTTTCTGAAAAAAAAATA 5099
```

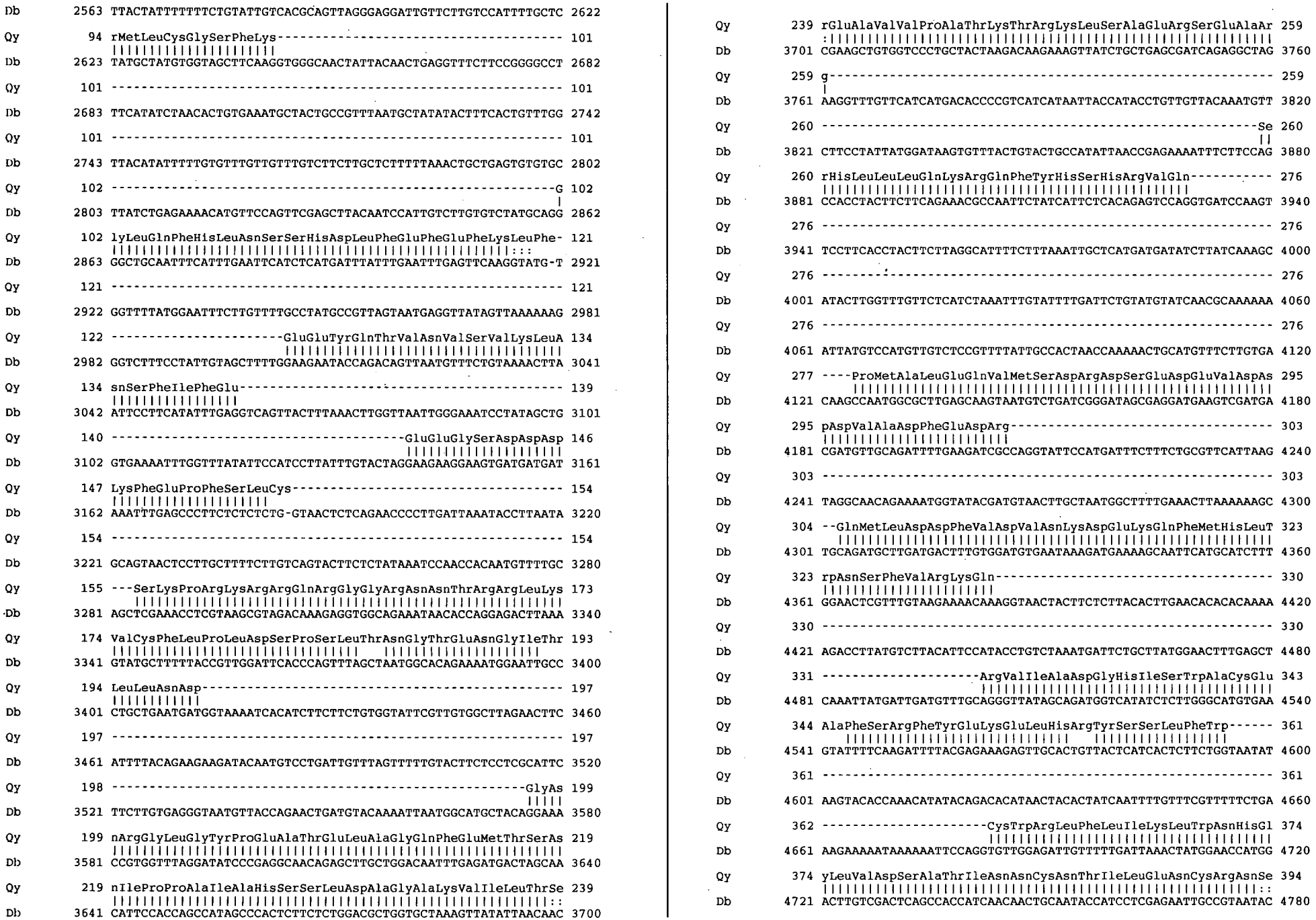
```
Qy       362 ------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAsp 377
                        ||||||||||||||||||||||||||||||||||||||||||||||||
Db      5100 AAAATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGAC 5159

Qy       378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5160 TCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACC 5219

Qy       398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5220 ACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAAT 5279

Qy       418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437
            |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5280 AACATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAAT 5339

Qy       438 AsnSerArgAspLysValIleLys 445
            |||||||||||||||||||||||
Db      5340 AACAGCAGAGACAAAGTAATTAAA 5363
```

```
RESULT 12
ATRPP5LE2/c
LOCUS       ATRPP5LE2              91660 bp    DNA     linear   PLN 22-DEC-2000
DEFINITION  Arabidopsis thaliana RPP5 disease resistance locus, main contig.
ACCESSION   AF180942 U97106
VERSION     AF180942.1  GI:6449044
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 76020 to 83095)
  AUTHORS   Parker,J.E., Coleman,M.J., Szabo,V., Frost,L.N., Schmidt,R., van
            der Biezen,E.A., Moores,T., Dean,C., Daniels,M.J. and Jones,J.D.
  TITLE     The Arabidopsis downy mildew resistance gene RPP5 shares similarity
            to the toll and interleukin-1 receptors with N and L6
  JOURNAL   Plant Cell 9 (6), 879-894 (1997)
  MEDLINE   97355983
   PUBMED   9212464
REFERENCE   2  (bases 1 to 91660)
  AUTHORS   Noel,L., Moores,T.L., van Der Biezen,E.A., Parniske,M.,
            Daniels,M.J., Parker,J.E. and Jones,J.D.
  TITLE     Pronounced intraspecific haplotype divergence at the RPP5 complex
            disease resistance locus of Arabidopsis
  JOURNAL   Plant Cell 11 (11), 2099-2112 (1999)
  MEDLINE   20025650
   PUBMED   10559437
REFERENCE   3  (bases 76020 to 83095)
  AUTHORS   Coleman,M.J., Parker,J.E., Szabo,V., van der Biezen,E.A.,
            Daniels,M.J. and Jones,J.D.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-1997) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich, Norfolk NR4 7UH, UK
REFERENCE   4  (bases 1 to 91660)
  AUTHORS   Noel,L., Moores,T., van der Biezen,E.A., Parniske,M., Daniels,M.J.,
            Parker,J.E. and Jones,J.D.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-1999) Sainsbury Laboratory, John Innes Centre,
            Colney Lane, Norwich, Norfolk NR4 7UH, United Kingdom
COMMENT     On Nov 18, 1999 this sequence version replaced gi:2109274.
FEATURES             Location/Qualifiers
     source          1. .91660
                     /organism="Arabidopsis thaliana"
                     /cultivar="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /map="C18a"
     gene            5354. .9673
                     /gene="La-I"
                     /pseudo
     CDS             join(5354. .5829,5977. .7057,7241. .7510,7610. .8071,
```

```
                8281. .9366,9458. .9673)
                /gene="La-I"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
repeat_region   15312. .20023
                /rpt_family="retroelement castor"
                /rpt_type=dispersed
LTR             15312. .15431
gene            15507. .19886
                /gene="encodes gag/pol"
                /pseudo
LTR             19904. .20023
gene            20048. .22736
                /gene="La-H"
                /pseudo
                /evidence=not_experimental
CDS             join(20048. .20440,20985. .22416,22533. .22736)
                /gene="La-H"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
gene            25823. .35584
                /gene="La-G"
                /note="interrupted by retroelement pollux1"
                /pseudo
                /evidence=not_experimental
CDS             join(25823. .26301,26428. .27343,32765. .32926,33111. .33407,
                33525. .33899,34108. .35235,35366. .35584)
                /gene="La-G"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
repeat_region   27344. .32762
                /rpt_family="retroelement pollux1"
                /rpt_type=dispersed
LTR             27344. .27791
gene            27937. .32166
                /gene="encodes gag/pol"
                /pseudo
LTR             32315. .32762
gene            36424. .39175
                /gene="La-F"
                /pseudo
                /evidence=not_experimental
CDS             join(36424. .37022,37167. .38250,38389. .38695,38855. .39175)
                /gene="La-F"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
gene            39222. .43288
                /gene="La-E"
                /pseudo
                /evidence=not_experimental
CDS             join(39222. .39754,39899. .40064,42749. .43288)
                /gene="La-E"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
gene            46918. .56304
                /gene="La-D"
                /note="interrupted by retroelement pollux2"
                /pseudo
                /evidence=not_experimental
CDS             join(46918. .47393,47558. .48635,48820. .48990,54412. .54573,
                54758. .55054,55172. .55546,55648. .55950,56080. .56304)
                /gene="La-D"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
repeat_region   48991. .54409
                /rpt_family="retroelement pollux2"
                /rpt_type=dispersed
LTR             48991. .49438
gene            49584. .53813
                /gene="encodes gag/pol"
LTR             53962. .54409
misc_feature    complement(58189. .58275)
                /note="similar to 5S ribosomal RNA gene repeat unit"
misc_feature    complement(58305. .58773)
                /note="similar to 5S ribosomal RNA gene repeat unit"
gene            59720. .65992
                /gene="La-C"
                /pseudo
                /evidence=not_experimental
CDS             join(59720. .60183,60265. .61340,62850. .63143,63293. .63703,
                63820. .64299,64540. .65646,65777. .65992)
                /gene="La-C"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
gene            68136. .72604
                /gene="La-B"
                /pseudo
                /evidence=not_experimental
CDS             join(68136. .68603,68739. .69793,69993. .70280,70398. .70870,
                71092. .72240,72368. .72604)
                /gene="La-B"
                /note="RPP5 homolog"
                /pseudo
                /codon_start=1
gene            <77312. .>82401
                /gene="RPP5"
                /evidence=experimental
mRNA            join(<77312. .77790,77917. .78991,79181. .79474,79581. .79994,
                80101. .80598,80811. .81842,81970. .82180,82319. .>82401)
                /gene="RPP5"
                /product="RPP5"
CDS             join(77312. .77790,77917. .78991,79181. .79474,79581. .79994,
                80101. .80598,80811. .81842,81970. .82180,82319. .82401)
                /gene="RPP5"
                /codon_start=1
                /product="downy mildew resistance protein RPP5"
                /protein_id="AAF08790.1"
                /db_xref="GI:6449046"
                /translation="MAASSSSGRRRYDVFPSFSGVDVRKTFLSHLLKALDGKSINTFI
                DHGIERSRTIAPELISAIREARISIVIFSKNYASSTWCLNELVEIHKCFNDLGQMVIP
                VFYDVDPSEVRKQTGEFGKVFEKTCEVSKDKQPGDQKQRWVQALTDIANIAGEDLLNG
                PNEAHMVEKISNDVSNKLITRSKCFDDFVGIEAHIEAIKSVLCLESKEARMVGIWGQS
                GIGKSTIGRALFSQLSSQFHHRAFLTYKSTSGSDVSGMKLSWQKELLSEILGQKDIKI
                EHFGVVEQRLNHKKVLILLDDVDNLEFLKTLVGKAEWFGSGSRIIVITQDRQLLKAHE
                IDLVYEVKLPSQGLALKMISQYAFGKDSPPDDFKELAFEVAELVGSLPLGLSVLGSSL
                KGRDKDEWVKMMPRLRNDSDDKIEETLRVGYDRLNKKNRELFKCIACFFNGFKVSNVK
                ELLEDDVGLTMLAEESLIRITPVGYIEMHNLLEKLGREIDRAKSKGNPGKRQFLTNFE
                DIREVLTEKTGTETLLGIRLPHPGYLTTRSFLIDEKSFKGMRNLQYLEIGYWSDGVLP
                QSLVYFPRKLKRLWWDNCPLKRLPSNFKAEYLVELRMVNSKLEKLWDGTQPLGSLKKM
                DLYNSYKLKEIPDLSLAINLEELNLEECESLETLPSSIQNAIKLRELNCWGGLLIDLK
                SLEGMCNLEYLSVPSWSSRECTQGIVYFPRKLKSVLWTNCPLKRLPSNFKAEYLVELI
                MEYSELEKLWDGTQSLGSLKEMNLRYSNNLKEIPDLSLAINLEELDLFGCVSLVTLPS
                SIQNATKLIYLDMSECENLESFPTVFNLKSLEYLDLTGCPNLRNFPAIKMGCAWTRLS
                RTRLFPEGRNEIVVEDCFWNKNLPAGLDYLDCLMRCMPCEFRSEQLTFLNVSGCKLEK
                LWEGIQSLGSLEEMDLSESENLKELPDLSKATNLKLLCLSGCKSLVTLPSTIGNLQNL
                RRLYMNRCTGLEVLPTDVNLSSLETLDLSGCSSLRTFPLISTNIVCLYLENTAIEEIP
                DLSKATKLESLILNNCKSLVTLPSTIGNLQNLRRLYMNRCTGLELLPTDVNLSSLETL
                DLSGCSSLRTFPLISTRIECLYLENTAIEEVPCCIEDFTRLTVLRMYCCQRLKNISPN
                IFRLTSLTLADFTDCRGVIKALSDATVVATMEDHVSCVPLSENIEYTCERFWDACSDY
                YSDDFEVNRNPIRLSTMTVNDVEFKFCCSITIKECGVRLLYVYQETEHNQQTTRSKKR
                MRMTSGTSEEDINLPYGQIVADTGLAALNTELSLGQGEASSSTSLEGEALCVDDYMIN
                EEQDEQIPILYPVYDIDDDMWRSLFFGDTDDDMWRSLYSAE"
BASE COUNT      28475 a  17266 c  17533 g  28386 t
ORIGIN

Alignment Scores:
Pred. No.:              8.02e-154       Length:         91660
Score:                  1798.50         Matches:        438
Percent Similarity:     43.38%          Conservative:   1
```

```
Best Local Similarity:  43.28%          Mismatches:    6
Query Match:            75.63%          Indels:        570
DB:                     8               Gaps:          13

US-09-890-220-2 (1-445) x ATRPP5LE2 (1-91660)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88263 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 88204

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      88203 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 88144

Qy         41 GlyAsn---------------------------------------------------------- 42
              ||||||
Db      88143 GGCAACGTATGATTTGCCTTCCTCTCTCATCATATTAGCTCAGTAATCTTTCATCTCCTG 88084

Qy         42 ---------------------------------------------------------------- 42

Db      88083 TGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGATTCATGGCGA 88024

Qy         43 ----------------------------------------------------ProSerPhe 45
                                                                   |||||||||
Db      88023 GTGTGTGCTTCTTTTGTCTCCTAAATTTGAACTTGTTGTTTGTTGTTGCAGCCATCGTTT 87964

Qy         46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer---------- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87963 CTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTA-TGTTTTCTT 87905

Qy         62 ---------------------------------------------------------------- 62

Db      87904 CTTGAATGTAGCTGCTACAGTGATATGTTATTTATCTTACTTCTAATATGGAAGCTGATG 87845

Qy         62 ---------------------------------------------------------------- 62

Db      87844 ACCTATTTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTGTTCATG 87785

Qy         63 ----------------------------------------ArgSerThrGlyMetVal 68
                                                      |||||||||||||||||
Db      87784 CTATAAACTTACATTTTATAAAATTGTGTTTTGGTTAGGTCAAGATCTACTGGGATGGTA 87725

Qy         69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------------- 83
              ||||||||||||||||||||||||||||||||||||||||||||
Db      87724 GTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGGTTAGTCTTTTTCT 87665

Qy         83 ---------------------------------------------------------------- 83

Db      87664 GTTCGTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTATTTACTA 87605

Qy         84 ----------------------ValArgGluAspCysSerCysProPheCysSerMetLe 96
                                     ||||||||||||||||||||||||||||||||||||
Db      87604 TTTTTTTCTGTATTGTCATGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCT 87545

Qy         96 uCysGlySerPheLys------------------------------------------- 101
              ||||||||||||||||
Db      87544 ATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCTTTCATA 87485

Qy        101 ---------------------------------------------------------------- 101

Db      87484 TCTAACACTGTGAAATGCTACTGCTGTTTCATGCTGTATACTTTCACTGTTTGGTTACAT 87425

Qy        101 ---------------------------------------------------------------- 101

Db      87424 ATTTTTGTGTTTGTTGTTTGTCTTCTCACTCTTTTCGAACTGCTGAGTGTGTGCTTATCT 87365

Qy        102 ----------------------------------------------------GlyLeuG 104
                                                                   |||||||
Db      87364 GAGAAAACATGTCCCAGATGGAGCTTACAACCAATTGTCTTGTGTCTATGCAGGGGCTGC 87305

Qy        104 lnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu---------- 120
              ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db      87304 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATGTGGTTTTA 87245

Qy        120 ---------------------------------------------------------------- 120

Db      87244 TGGAAATTCTTGATTTGCTATGCCTTTATTAATGAGGTTATAGTTAAAAAAGGGTCTTTC 87185

Qy        121 -----------Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
                          ||| ||||||||||||||||||||||||||||||||||||||||||||
Db      87184 CTATTGTAGCTTTCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTC 87125

Qy        137 IlePheGluGlu------------------------------------------------- 140
              |||||||||
Db      87124 ATATTTGAAGT-CAGTAACTTTAAACTTGGTTAATTGGGAAATCCGATAGCTGGTGAAAA 87066

Qy        141 -----------------------------------GluGlySerAspAspAspLysPhe 148
                                                  |||||||||||||||||||||||
Db      87065 TTTTGTTTATATTCCATCCTTATTTGTACTAGGAAAGAAGGAAGTGATGACGATAAATTT 87006

Qy        149 GluProPheSerLeu------------------------------------------- 153
              ||||||||||||||||
Db      87005 GAGCCCTTCTCTCTCTGGTAACCCTCAGAACCCTTCCGATTAAATACCTTAATAGCAGTA 86946

Qy        154 ----------------------------------------------------CysSer 155
                                                                   |||
Db      86945 ACTCCTTGCTTCTCTTGTCAGTACATCTCTGTAAATCCAACCATAATGTTTTGCAGCTCG 86886

Qy        156 LysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCys 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86885 AAACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGC 86826

Qy        176 PheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86825 TTTTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTT 86766

Qy        196 AsnAsp------------------------------------------------------ 197
              ||||||
Db      86765 AATGATGGTAAAATCATATCTTCTTCTGTGCGTTCCTTGTGGCTTAGAACTTCATATTAC 86706

Qy        197 ---------------------------------------------------------------- 197

Db      86705 AGAAGAAGATACAATGGCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTCTTCTTGC 86646

Qy        198 ----------------------------------------------------GlyAsnArgGly 201
                                                                   ||||||||||||
Db      86645 GAGGGTATTGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAACCGTGGT 86586

Qy        202 LeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIlePro 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86585 TTAGGATATCCCGAGGCAACAGAACTTGCTGGACAATTTGAGATGACCAGCAACATTCCA 86526

Qy        222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAla 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86525 CCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCT 86466

Qy        242 ValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg------ 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86465 GTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGGTTT 86406

Qy        259 ---------------------------------------------------------------- 259

Db      86405 GTTCATCATGACACCCCGTCATCATAATTACCATTCCTGTTGTTACAAATGTTCTTCCTA 86346

Qy        260 ----------------------------------------------------SerHisLeu 262
                                                                   ||||||||
Db      86345 TTATGGATAAGTGTTTATAGTACTGCCATATTAACCGAGAAAATTTCTTCCAGCCACCTA 86286

Qy        263 LeuLeuGlnLysArgGlnPheTyrHisSerHisArgVal---------------------- 275
              ||||||||||||||||||||||||||||||||||||.|
Db      86285 CTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGTTCCTTCA 86226

Qy        275 ---------------------------------------------------------------- 275

Db      86225 CCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGCATACTTG 86166
```

```
Qy        275 ------------------------------------------------------------ 275

Db      86165 GTTTGTTCTCATCCAAATTTGTATTTTGATCTGTATGTATCAACGCAAAATAGTTATGTC 86106

Qy        276 ----------------------------------------------------Gln-ProMe 278
                                                                   ||| ||||| 
Db      86105 CATGTTGTCTCCGTTTTATTGCCACTAACCAAAAAATGCATGTTTCTGTGACAAGCCAAT 86046

Qy        278 tAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAl 298
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      86045 GGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGC 85986

Qy        298 aAspPheGluAspArg---------------------------------------------- 303
              |||||||||||||||  
Db      85985 AGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAATAGACAACA 85926

Qy        304 ----------------------------------------------------GlnMetL 306
                                                                   ||||||| 
Db      85925 GAAAATGGTATATGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGCTGCAGATGC 85866

Qy        306 euAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerP 326
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      85865 TTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGT 85806

Qy        326 heValArgLysGlnArgVal------------------------------------------ 332
              ||||||||||||||||  
Db      85805 TTGTAAGAAAACAAAGGTA-ACTACTTCTCTTACACATGAACAGACACAAAAAGACCTTA 85747

Qy        332 ------------------------------------------------------------ 332

Db      85746 TGTCTTACATTCCATACCTGTCTAAATGATTTTGCTTATGGAACTTTGAGCTCAATTATG 85687

Qy        333 --------------------IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheS 346
                                   |||||||||||||||||||||||||||||||||||||| 
Db      85686 ATTGTTGATGTTTCAGGGTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTT 85627

Qy        346 erArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------------- 361
              ||||||||||||||||||||||||||||||||||||||||||||| 
Db      85626 CAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTAATATAAGTACA 85567

Qy        361 ------------------------------------------------------------ 361

Db      85566 CCAAACATATACAGACACATAACTACACTATCAATCTTGTTTCGTTTTCTGAAAAAAAAA 85507

Qy        362 --------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValA 377
                           |||||||||||||||||||||||||||||||||||||||||||| 
Db      85506 TAAAAATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCG 85447

Qy        377 spSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspT 397
              |||||||||||||||||||||||||||||||||||||||   |||||||||||||||| 
Db      85446 ACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGGCGTAATAGCTCAGACA 85387

Qy        397 hrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnA 417
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      85386 CCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACA 85327

Qy        417 snAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspA 437
              ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||| 
Db      85326 ATAACATTGTGAGTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACA 85267

Qy        437 snAsnSerArgAspLysValIleLys 445
              ||||||||||||||||||||||||| 
Db      85266 ATAACAGCAGAGACAAAGTAATTAAA 85241

RESULT 13
AX032893
LOCUS       AX032893               5895 bp    DNA     linear   PAT 21-SEP-2000
DEFINITION  Sequence 6 from Patent WO0044918.
ACCESSION   AX032893
VERSION     AX032893.1  GI:10279828
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 5895)
  AUTHORS   Dean,C. and Gendall,A.
  TITLE     Methods and means for modification of plant characteristics using
            the vernalization gene vrn2
  JOURNAL   Patent: WO 0044918-A 6 03-AUG-2000;
            DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
            (GB)
FEATURES             Location/Qualifiers
     source          1. .5895
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
BASE COUNT     1742 a   1027 c   1070 g   2056 t
ORIGIN

Alignment Scores:
Pred. No.:                1.44e-142      Length:          5895
Score:                    1659.00        Matches:         424
Percent Similarity:       42.25%         Conservative:    4
Best Local Similarity:    41.86%         Mismatches:      12
Query Match:              69.76%         Indels:          577
DB:                       6              Gaps:            14

US-09-890-220-2 (1-445) x AX032893 (1-5895)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1905 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 1964

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1965 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2024

Qy         41 GlyAsn-------------------------------------------------------- 42
              ||| 
Db       2025 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 2083

Qy         42.------------------------------------------------------------ 42

Db       2084 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 2143

Qy         43 ------------------------------------------------------------Pro 43
                                                                           ||| 
Db       2144 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 2203

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||    
Db       2204 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 2262

Qy         62 ------------------------------------------------------------ 62

Db       2263 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 2322

Qy         62 ------------------------------------------------------------ 62

Db       2323 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 2382

Qy         63 ----------------------------------------------------ArgSerThrGly 66
                                                                   |||||||||| 
Db       2383 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 2442

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              |||||||||||||||||||||||||||||||||||||:::|||||| 
Db       2443 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 2502

Qy         83 ------------------------------------------------------------ 83

Db       2503 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 2562

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           |||||||||||||||||||||||||||||| 
```



```
Qy        394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
              :|||         |||    ||||||||||||||||||||    ||||||||||||||||
Db       4781 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 4834

Qy        414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       4835 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 4894

Qy        434 nLysAspAsnAsnSerArgAspLys 442
              |||||||||||||||||||||||||
Db       4895 CAAGGACAATAACAGCAGAGACAAG 4919

RESULT 14
ATCHRIV45
LOCUS       ATCHRIV45              199548 bp    DNA     linear   PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 45.
ACCESSION   AL161545
VERSION     AL161545.2  GI:7268431
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 199548)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV44 at the 5' end and an
            overlap with ATCHRIV46 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .199548
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     gene            6008. .10944
                     /gene="AT4g16810"
     gene            complement(join(6008. .6649,6855. .7090,7222. .7258,
                     7337. .7536,7629. .7772,7877. .7913,8744. .8853,9020. .9108,
                     9487. .9576,9704. .9786,9891. .9980,10178. .10229,
                     10883. .10944))
                     /gene="AT4g16810"
     CDS             complement(join(6008. .6649,6855. .7090,7222. .7258,
                     7337. .7536,7629. .7772,7877. .7913,8744. .8853,9020. .9108,
                     9487. .9576,9704. .9786,9891. .9980,10178. .10229,
                     10883. .10944))
                     /gene="AT4g16810"
                     /note="similarity to centromere protein E, human"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB80952.1"
                     /db_xref="GI:7268432"
                     /translation="MDPIKLTTEAKVPAKRSKATSHYLPLHKRQFYHSRTGQPLSLEQ
                     VMSDRDSENDVDKNDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAH
                     IPWACEAFSRLHLQELRSNLSLDLGNTETTETAPVATEATTHGEQDGIDVGEQEAVVV
                     EVVVAEAEVVVVVKAEAQAVVETSSSDLRWAGEDGYSDSTSPISEPPDITNWFPSYVY
                     ESLPLSSSINEIEIEVDVVNKANPGGVIDESLPSEPPDLGKWFSSYVYESPLLDTSDG
                     LELSVPGESECVKETQTENESPKIEGNDVCPRLLEQELVSSTKVTDFSESQSVLSEPP
                     DLRNWFSSYEYQSPQLSDIQEFGFLYSEKDELIIEESDTEEGISSGIFRKIKSRQETV
                     GLGRLDSNDYKENIATDTAKEVSLDNAVSDQKMEKRSSVRLLNASKKDVKQESSFKQE
                     PLLCELQEEARFSPRVSRYNPKLKSPSKNDTSLHELRPIHIQESISMNSNRQKSPPIE
                     QESYDKENVHGQSNQTGFVTMKKASFREARDQCSLKKPNTEVLVKCSSSKELKTKAGE
                     DNTEEREKKRRVLGEMSNQLSSVAKEIAGKWRCPQKNKRKIVPPLKQLRLDAWIHKV"
     exon            complement(6008. .6649)
                     /gene="AT4g16810"
                     /number=1
     intron          complement(6650. .6854)
                     /gene="AT4g16810"
                     /number=1
     exon            complement(6855. .7090)
                     /gene="AT4g16810"
                     /number=2
     intron          complement(7091. .7221)
                     /gene="AT4g16810"
                     /number=2
     exon            complement(7222. .7258)
                     /gene="AT4g16810"
                     /number=3
     intron          complement(7259. .7336)
                     /gene="AT4g16810"
                     /number=3
     exon            complement(7337. .7536)
                     /gene="AT4g16810"
                     /number=4
     intron          complement(7537. .7628)
                     /gene="AT4g16810"
                     /number=4
     exon            complement(7629. .7772)
                     /gene="AT4g16810"
                     /number=5
     intron          complement(7773. .7876)
                     /gene="AT4g16810"
                     /number=5
     exon            complement(7877. .7913)
                     /gene="AT4g16810"
                     /number=6
     intron          complement(7914. .8743)
                     /gene="AT4g16810"
                     /number=6
     exon            complement(8744. .8853)
                     /gene="AT4g16810"
                     /number=7
     intron          complement(8854. .9019)
                     /gene="AT4g16810"
                     /number=7
     exon            complement(9020. .9108)
                     /gene="AT4g16810"
                     /number=8
     intron          complement(9109. .9486)
                     /gene="AT4g16810"
                     /number=8
     exon            complement(9487. .9576)
                     /gene="AT4g16810"
                     /number=9
     intron          complement(9577. .9703)
                     /gene="AT4g16810"
                     /number=9
     exon            complement(9704. .9786)
                     /gene="AT4g16810"
                     /number=10
     intron          complement(9787. .9890)
                     /gene="AT4g16810"
                     /number=10
     exon            complement(9891. .9980)
                     /gene="AT4g16810"
                     /number=11
     intron          complement(9981. .10177)
                     /gene="AT4g16810"
                     /number=11
     exon            complement(10178. .10229)
                     /gene="AT4g16810"
                     /number=12
     intron          complement(10230. .10882)
                     /gene="AT4g16810"
                     /number=12
     exon            complement(10883. .10944)
                     /gene="AT4g16810"
```

```
                /number=13
     gene       16254. .19247
                /gene="AT4g16820"
     CDS        join(16254. .17790,18979. .19247)
                /gene="AT4g16820"
                /note="similarity to triacylglycerol lipase (EC 3.1.1.3),
                Rhizopus niveus
                Contains Lipases, serine active site AA321-330
                contains EST gb:AI993282.1, T76589, AI994595.1, R30062,
                Z32589"
                /codon_start=1
                /product="triacylglycerol lipase like protein"
                /protein_id="CAB80953.1"
                /db_xref="GI:7268433"
                /translation="MQTLTPNADIFHAKRRRFTCNTHTSTLIPTKPLSVSPARKTNKE
                HLRNLENVLRTSSNSIDHIENVTSRQEKTTKNTSTSSLLGGLNLARIWPQMKAAVDEM
                SPKNLKRLQRLLSKSSEERSPKSKLGSKWRELHGLNNWAGLLDPLDENLRRELVRYGE
                FVQAAYHAFHSDPEGSPRHVALPDGSFKVTKSLYATSSVRLPKWIDDVAPDLRWMTKQ
                TSWVGYVAVCDDPREIRRMGRREIVIALRGTATLLEWSENFRPNLVSMPEPKPDQSDP
                TRPKVECGFNSLYTTGDQHAPSLAESLVGEISRLVELYAGEELSISVTGHSLGAAIAL
                LAADDIAERVPHAPPVAVFSFGGPRVGNREFADRLDSKGVKVLRVVNSQDVVTKVPGI
                FADNDKQGQSRNNGRSPGGIMEMVERNNPWAYSHVGAELRVDMKMSPYLKPNADVACC
                HDLEAYLHLVDGFLASNCPFRANAKRSLRKLLDEQRSNVKVLYTGKSLRLNRTFHDNG
                DVLPRDTGQGDNPIMKYKLLFRYAVGLSILIITTVFLPIPNFHTFYFHFLIFISLSFL
                SLTNQQVLLSTHFVSFILTFSLDLLIIPSMAWSSH"
     exon       16254. .17790
                /gene="AT4g16820"
                /number=1
     intron     17791. .18978
                /gene="AT4g16820"
                /number=1
     exon       18979. .19247
                /gene="AT4g16820"
                /number=2
     gene       19312. .20999
                /gene="AT4g16830"
     CDS        join(19312. .19483,19652. .19958,20041. .20106,20210. .20336,
                20437. .20592,20679. .20729,20811. .20999)
                /gene="AT4g16830"
                /note="cognate cDNA CDNA4KMY is matching at position
                1428395-1430523
                similarity to nuclear antigen - Bos taurus
                contains EST gb:AI994595.1, AI993282.1, R30062, T42546,
                Z32590, Z32589, T76589, AA395095"
                /codon_start=1
                /product="nuclear antigen homolog"
                /protein_id="CAB80954.1"
                /db_xref="GI:7268434"
                /translation="MATLNPFDLLDDDAEDPSQLAVAIEKIDKSKKSGQVSSLPAKSA
                PKLPSKPLPPAQAVREARSDAPRGGGGRGGFNRGRGGYNRDDGNNGYSGGYTKPSGEG
                DVSKSSYERRGGGGAPRGSFRGEGGGPGGGRRGGFSNEGGDGERPRRAFERRSGTGRG
                SDFKRDGSGRGNWGTPGEEIAAETEAVAGVETEKDVGEKPAVDDVAADANKEDTVVEE
                KEPEDKEMTLDEYEKILEEKKKALQSLTTSERKVDTKVFESMQQLSNKKSNDEIFIKL
                GSDKDKRKDDKEEKAKKAVSINEFLKPAEGGNYYRGGRGGRGRGGRGRGGVSSGESGG
                YRNEAAPAIGDAAQFPSLGGK"
     exon       19312. .19483
                /gene="AT4g16830"
                /number=1
     intron     19484. .19651
                /gene="AT4g16830"
                /number=1
     exon       19652. .19958
                /gene="AT4g16830"
                /number=2
     intron     19959. .20040
                /gene="AT4g16830"
                /number=2
     exon       20041. .20106

Alignment Scores:
Pred. No.:              1.65e-140       Length:         199548
Score:                  1659.00         Matches:        424
Percent Similarity:     42.25%          Conservative:   4
Best Local Similarity:  41.86%          Mismatches:     12
Query Match:            69.76%          Indels:         577
DB:                     8               Gaps:           14

US-09-890-220-2 (1-445) x ATCHRIV45 (1-199548)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25399 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 25458

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25459 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 25518

Qy         41 GlyAsn---------------------------------------------------------- 42
              |||
Db      25519 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 25577

Qy         42 ---------------------------------------------------------------- 42

Db      25578 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 25637

Qy         43 ----------------------------------------------------------Pro 43
                                                                         |||
Db      25638 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 25697

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25698 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 25756

Qy         62 ---------------------------------------------------------------- 62

Db      25757 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 25816

Qy         62 ---------------------------------------------------------------- 62

Db      25817 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 25876

Qy         63 ----------------------------------------------------ArgSerThrGly 66
                                                                   ||||||||||||
Db      25877 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 25936

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              ||||||||||||||||||||||||||||||||||||||:::|||||||
Db      25937 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 25996

Qy         83 ---------------------------------------------------------------- 83

Db      25997 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 26056

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           ||||||||||||||||||||||||||||||
Db      26057 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 26116

Qy         94 rMetLeuCysGlySerPheLys---------------------------------------- 101
              ||||||||||||||||||||
Db      26117 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 26176

Qy        101 ---------------------------------------------------------------- 101

Db      26177 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 26236

Qy        101 ---------------------------------------------------------------- 101

Db      26237 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 26296

Qy        102 ----------------------------------------------------------G 102
                                                                         |
Db      26297 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 26356

Qy        102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db      26357 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 26415
```

```
Qy      121 ------------------------------------------------------------ 121
Db    26416 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 26475
Qy      122 ----------------------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
                                   ||||||||||||||||||||||||||||||||||||
Db    26476 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 26535
Qy      134 snSerPheIlePheGlu---------------------------------------- 139
              |||||||||||||||||
Db    26536 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 26595
Qy      140 ---------------------------------------GluGluGlySerAspAspAsp 146
                                                    ||||||||||||||||||||
Db    26596 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 26655
Qy      147 LysPheGluProPheSerLeuCys---------------------------------- 154
              ||||||||||||||||||||
Db    26656 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 26714
Qy      154 ------------------------------------------------------------ 154
Db    26715 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 26774
Qy      155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    26775 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 26834
Qy      174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
              |||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db    26835 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 26894
Qy      194 LeuLeuAsnAsp------------------------------------------------ 197
              |||||||||||
Db    26895 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 26954
Qy      197 ------------------------------------------------------------ 197
Db    26955 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 27014
Qy      198 ------------------------------------------------------GlyAs 199
                                                                     |||||
Db    27015 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 27074
Qy      199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27075 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 27134
Qy      219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db    27135 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 27194
Qy      239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27195 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 27254
Qy      259 g------------------------------------------------------------ 259
              |
Db    27255 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 27314
Qy      260 ------------------------------------------------------------Se 260
                                                                          ||
Db    27315 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 27374
Qy      260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----------- 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27375 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 27434
Qy      276 ------------------------------------------------------------ 276
Db    27435 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 27494

Qy      276 ------------------------------------------------------------ 276
Db    27495 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 27554
Qy      276 ------------------------------------------------------------ 276
Db    27555 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 27614
Qy      277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27615 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 27674
Qy      295 pAspValAlaAspPheGluAspArg------------------------------------ 303
              |||||||||||||||||||||||
Db    27675 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 27734
Qy      303 ------------------------------------------------------------ 303
Db    27735 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 27794
Qy      304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27795 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 27854
Qy      323 rpAsnSerPheValArgLysGln---------------------------------- 330
              |||||||||||||||||||||
Db    27855 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 27914
Qy      330 ------------------------------------------------------------ 330
Db    27915 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 27974
Qy      331 ----------------------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
                                   ||||||||||||||||||||||||||||||||||||||
Db    27975 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 28034
Qy      344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
              ||||||||||||||||||||||||||||||   ||||||||||||||||||
Db    28035 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 28094
Qy      361 ------------------------------------------------------------ 361
Db    28095 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 28154
Qy      362 ----------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
                                   ||||||||||||||||||||||||||||||||||||||
Db    28155 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 28214
Qy      374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db    28215 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 28274
Qy      394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
              :|||         |||   ||||||||||||||||||||   ||||||||||||||||||
Db    28275 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 28328
Qy      414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db    28329 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 28388
Qy      434 nLysAspAsnAsnSerArgAspLys 442
              |||||||||||||||||||||||
Db    28389 CAAGGACAATAACAGCAGAGACAAG 28413
```

```
RESULT 15
ATFCA7
LOCUS       ATFCA7                201471 bp    DNA     linear   PLN 29-JUN-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment
            No. 7.
ACCESSION   Z97342
VERSION     Z97342.2  GI:5302800
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
```

```
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 201471)
  AUTHORS   Bevan,M., Stiekema,W., Murphy,G., Wambutt,R., Pohl,T., Terryn,N.,
            Kreis,M., Kavanagh,T., Entian,K.D., Rieger,M., James,R.,
            Puigdomenech,P., Hatzopoulos,P., Obermaier,B., Duesterhoft,A.,
            Jones,J., Palme,K., Ansorge,W., Delseny,M., Bancroft,I.,
            Mewes,H.W., Schueller,C. and Chalwatzis,N.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201471)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     On Jun 30, 1999 this sequence version replaced gi:2245031.
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATFCA6 at the 5' end and an
            overlap with ATFCA8 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .201471
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     source          1. .8956
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB5K23"
     source          21108. .37844
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB6C5"
     source          37120. .39085
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid EW16B10.IB4A23"
     source          37904. .128469
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="BAC IGF5D3"
     source          38368. .63818
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC33M3"
     source          106243. .128469
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC16N19"
     source          113141. .132730
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC20I2"
     source          129847. .157017
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid CC1B24"
     source          152677. .190349
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid G16599"
     source          185303. .201471
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="cosmid G11319"
     gene            12153. .13466
                     /gene="dl4415w"
     CDS             join(12153. .12312,12416. .12714,12983. .13062,13151. .13466)
                     /gene="dl4415w"
                     /note="contains EST gb:
                     1216777;935272;906153;947895;315626"
                     /codon_start=1
                     /product="DNA-binding homeotic protein Athb-2"
                     /protein_id="CAB10452.1"
                     /db_xref="GI:2245033"
                     /db_xref="SWISS-PROT:Q05466"
                     /translation="MMFEKDDLGLSLGLNFPKKQINLKSNPSVSVTPSSSSFGLFRRS
                     SWNESFTSSVPNSDSSQKETRTFIRGIDVNRPPSTAEYGDEDAGVSSPNSTVSSSTGK
                     RSEREEDTDPQGSRGISDDEDGDNSRKKLRLSKDQSAILEETFKDHSTLNPKQKQALA
                     KQLGLRARQVEVWFQNRRARTKLKQTEVDCEFLRRCCENLTEENRRLQKEVTELRALK
                     LSPQFYMHMSPPTTLTMCPSCEHVSVPPPQPQAATSAHHRSLPVNAWAPATRISHGLT
                     FDALRPRS"
     exon            12153. .12312
                     /gene="dl4415w"
                     /number=1
     intron          12313. .12415
                     /gene="dl4415w"
                     /number=1
     exon            12416. .12714
                     /gene="dl4415w"
                     /number=2
     intron          12715. .12982
                     /gene="dl4415w"
                     /number=2
     exon            12983. .13062
                     /gene="dl4415w"
                     /number=3
     intron          13063. .13150
                     /gene="dl4415w"
                     /number=3
     exon            13151. .13466
                     /gene="dl4415w"
                     /number=4
     gene            14609. .16030
                     /gene="dl4420c"
     gene            complement(14609. .16030)
                     /gene="dl4420c"
     CDS             complement(14609. .16030)
                     /gene="dl4420c"
                     /note="sequence differences between cDNA and gDNA
                     similarity to hydroxyproline-rich glycoprotein (clone
                     Hyp3.6) - kidney bean (fragment)"
                     /codon_start=1
                     /product="glycoprotein homolog"
                     /protein_id="CAB46042.1"
                     /db_xref="GI:5302801"
                     /translation="MVEARSLKKPIQLGNKEDQNPRKFYSRFIFKALILTVLCAVVPV
                     FLSQTPELANQTRLLELLHLVFVGIAVSYGLFSRRNYDGGGGGGTSNSDHNKADHSNN
                     NSHSYVPKILEVSSVFNVGHESESEPSDDSSGDQRKFQTWKNKYHMKIPEVETRFVDR
                     VSSENREKPLLLPVRSLNYSRVSDSSGDNSGRWEKVRSKRELLKTLGDDNSDVLPSPI
                     PWRSRSSSSSSSSSKEVESLPSVKNLTTVESQPLIKNLTPPSSFSSPRKSNPIPNLAS
                     EFHPSPPPPPPPPPPLPAFYNSSSRKDHPGIYRVERRESSVHKTKFAGGEFHPPPPPP
                     PPPPVEYYKSPPTKFRLSNERRKSSEQKMKRNAPKKVWWSDPIVESKEQDTEKNDQRS
                     NLGSKAVEESENGEQRRGENEIHDEVEKKIVEEEGVSEINNGSDVDKKADEFIAKFRE
                     QIRLQRIESIKRSTNKISANSSR"
     exon            complement(14609. .16030)
                     /gene="dl4420c"
                     /number=1
     repeat_region   16148. .16170
                     /note="23bp T tandem repeat"
     misc_feature    17691. .19115
                     /note="CDNA1KMY"
     gene            17793. .19171
                     /gene="dl4425c"
     gene            complement(join(17793. .18046,18147. .18250,18348. .18399,
                     18481. .18571,18886. .18960,19058. .19171))
                     /gene="dl4425c"
     CDS             complement(join(17793. .18046,18147. .18250,18348. .18399,
                     18481. .18571,18886. .18960,19058. .19171))
```

```
                  /gene="dl4425c"
                  /note="strong similarity to enoyl-CoA hydratase (EC
                  4.2.1.17), AU-specific RNA-binding protein, human"
                  /codon_start=1
                  /product="enoyl-CoA hydratase"
                  /protein_id="CAB10453.1"
                  /db_xref="GI:2245034"
                  /db_xref="SPTREMBL:O23520"
                  /translation="MLKSLQNAFESIHQDNSARVVMIRSLVPGVFCAGADLKERRTMS
                  PSEVHTYVNSLRYMFSFIEALSIPTIAAIEGAALGGGLEMALACDLRICGENAVFGLP
                  ETGLAIIPGAGGTQRLSRLVGRSVSKELIFTGRKIDAIEAANKGLVNICVTAGEAHEK
                  AIEMAQQINEKGPLAIKMAKKAIDEGIETNMASGLEVEEMCYQKLLNTQDRLEGLAAF
                  AEKRKPLYTGN"
     exon         complement(17793. .18046)
                  /gene="dl4425c"
                  /number=1
     intron       complement(18047. .18146)
                  /gene="dl4425c"
                  /number=1
     exon         complement(18147. .18250)
                  /gene="dl4425c"
                  /number=2
     intron       complement(18251. .18347)
                  /gene="dl4425c"
                  /number=2
     exon         complement(18348. .18399)
                  /gene="dl4425c"
                  /number=3
     intron       complement(18400. .18480)
                  /gene="dl4425c"
                  /number=3
     exon         complement(18481. .18571)
                  /gene="dl4425c"
                  /number=4
     intron       complement(18572. .18885)
                  /gene="dl4425c"
                  /number=4
     exon         complement(18886. .18960)
                  /gene="dl4425c"
                  /number=5
     intron       complement(18961. .19057)
                  /gene="dl4425c"
                  /number=5
     exon         complement(19058. .19171)
                  /gene="dl4425c"
                  /number=6
     gene         20179. .25115

Alignment Scores:
Pred. No.:              1.67e-140     Length:          201471
Score:                  1659.00       Matches:         424
Percent Similarity:     42.25%        Conservative:    4
Best Local Similarity:  41.86%        Mismatches:      12
Query Match:            69.76%        Indels:          577
DB:                     8             Gaps:            14

US-09-890-220-2 (1-445) x ATFCA7 (1-201471)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39570 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 39629

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39630 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 39689

Qy         41 GlyAsn---------------------------------------------------- 42
              |||
Db      39690 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 39748

Qy         42 ---------------------------------------------------------- 42

Db      39749 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 39808

Qy         43 ----------------------------------------------------------Pro 43
                                                                          |||
Db      39809 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 39868

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39869 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 39927

Qy         62 ---------------------------------------------------------- 62

Db      39928 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 39987

Qy         62 ---------------------------------------------------------- 62

Db      39988 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 40047

Qy         63 ----------------------------------------------------ArgSerThrGly 66
                                                                   |||||||||||
Db      40048 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 40107

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              ||||||||||||||||||||||||||||||||||||||||:::|||||||
Db      40108 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 40167

Qy         83 ---------------------------------------------------------- 83

Db      40168 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 40227

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           ||||||||||||||||||||||||||||||
Db      40228 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 40287

Qy         94 rMetLeuCysGlySerPheLys---------------------------------- 101
              ||||||||||||||||||||
Db      40288 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 40347

Qy        101 ---------------------------------------------------------- 101

Db      40348 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 40407

Qy        101 ---------------------------------------------------------- 101

Db      40408 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 40467

Qy        102 ----------------------------------------------------------G 102
                                                                           |
Db      40468 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 40527

Qy        102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db      40528 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 40586

Qy        121 ---------------------------------------------------------- 121

Db      40587 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 40646

Qy        122 ----------------------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
                                     ||||||||||||||||||||||||||||||||||||
Db      40647 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 40706

Qy        134 snSerPheIlePheGlu------------------------------------------ 139
              ||||||||||||||||
Db      40707 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 40766

Qy        140 ----------------------------------------GluGluGlySerAspAspAsp 146
                                                       ||||||||||||||||||||
Db      40767 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 40826

Qy        147 LysPheGluProPheSerLeuCys---------------------------------- 154
              ||||||||||||||||||||
Db      40827 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 40885
```

```
Qy       154 ------------------------------------------------------------ 154

Db     40886 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 40945

Qy       155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40946 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 41005

Qy       174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
                 |||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db     41006 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 41065

Qy       194 LeuLeuAsnAsp---------------------------------------------- 197
                 ||||||||||||
Db     41066 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 41125

Qy       197 ------------------------------------------------------------ 197

Db     41126 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 41185

Qy       198 ------------------------------------------------------GlyAs 199
                                                                     |||||
Db     41186 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 41245

Qy       199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41246 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 41305

Qy       219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db     41306 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 41365

Qy       239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
                 :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41366 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 41425

Qy       259 g------------------------------------------------------------ 259
                 |
Db     41426 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 41485

Qy       260 ------------------------------------------------------------Se 260
                                                                           ||
Db     41486 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 41545

Qy       260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln---------- 276
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41546 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 41605

Qy       276 ------------------------------------------------------------ 276

Db     41606 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 41665

Qy       276 ------------------------------------------------------------ 276

Db     41666 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 41725

Qy       276 ------------------------------------------------------------ 276

Db     41726 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 41785

Qy       277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41786 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 41845

Qy       295 pAspValAlaAspPheGluAspArg------------------------------------ 303
                 ||||||||||||||||||||||
Db     41846 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 41905

Qy       303 ------------------------------------------------------------ 303

Db     41906 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 41965

Qy       304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323

                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41966 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 42025

Qy       323 rpAsnSerPheValArgLysGln---------------------------------- 330
                 ||||||||||||||||||||
Db     42026 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 42085

Qy       330 ------------------------------------------------------------ 330

Db     42086 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 42145

Qy       331 ----------------------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
                                      ||||||||||||||||||||||||||||||||||||
Db     42146 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 42205

Qy       344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
                 ||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     42206 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 42265

Qy       361 ------------------------------------------------------------ 361

Db     42266 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 42325

Qy       362 ----------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
                                    ||||||||||||||||||||||||||||||||||||
Db     42326 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 42385

Qy       374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db     42386 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 42445

Qy       394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
                 :|||         |||   ||||||||||||||||||||   ||||||||||||||||
Db     42446 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 42499

Qy       414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
                 ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db     42500 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 42559

Qy       434 nLysAspAsnAsnSerArgAspLys 442
                 ||||||||||||||||||||||||
Db     42560 CAAGGACAATAACAGCAGAGACAAG 42584
```

Search completed: June 20, 2003, 00:55:20
Job time : 3851 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 22:08:48 ; Search time 330 Seconds
(without alignments)
3036.786 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144850_16438/app_query.fasta_1.583
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_263_@runat_13062003_144850_16438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2378	100.0	1722	21	AAA63670	Nucleotide sequenc
	2	2378	100.0	1722	21	AAA47751	VRN2 nucleic acid.
	3	2285	96.1	1715	21	AAA63671	Nucleotide sequenc
	4	2285	96.1	1715	21	AAA47753	VRN2 nucleic acid.
	5	2258.5	95.0	1737	21	AAA47755	Aberrantly spliced
	6	1923	80.9	1497	21	AAC51581	Arabidopsis thalia
	7	1813.5	76.3	6338	21	AAA47752	VRN2 nucleic acid.
	8	1659	69.8	5895	21	AAA47754	VRN2 nucleic acid.
	9	923	38.8	2280	21	AAZ36947	cDNA encoding a MP
	10	906.5	38.1	2248	21	AAZ36953	cDNA encoding an O
	11	451.5	19.0	558	21	AAA47758	At Hyp 2245035 (AT
	12	407.5	17.1	5580	21	AAZ36948	DNA encoding a MPC
	13	368.5	15.5	325	21	AAA47757	AI163743 Expressed
	14	340.5	14.3	640	21	AAC44147	Arabidopsis thalia
	15	327	13.8	319	21	AAC51582	Arabidopsis thalia
	16	324	13.6	525	21	AAC54919	Arabidopsis thalia
	17	311	13.1	463	21	AAA47756	Rice expressed seq
	18	225	9.5	3545	22	AAS25965	Human cDNA encodin
	19	221	9.3	4441	21	AAA47759	KIAA0160 cDNA. Ho
	20	221	9.3	4441	24	ABL69961	Pancreas cancer re
	21	219	9.2	3572	23	ABL29637	Drosophila melanog
	22	208	8.7	7021	23	ABL29636	Drosophila melanog
	23	196.5	8.3	1216	21	AAZ36962	DNA fragment encod
	24	157.5	6.6	1257	24	AAS62641	cDNA sequence #428
	25	153	6.4	2481	22	AAS25945	Human cDNA encodin
	26	131.5	5.5	3738	21	AAA70178	Plasmodium falcipa
	27	128.5	5.4	5454	21	AAA70236	Plasmodium falcipa
	28	127.5	5.4	3300	21	AAZ55699	DNA encoding yeast
	29	127	5.3	2466	23	ABL27385	Drosophila melanog
	30	127	5.3	8045	21	AAA65171	Neurospora crassa
c	31	126	5.3	646	24	ABQ46952	Oligonucleotide fo
	32	126	5.3	646	24	ABQ46953	Oligonucleotide fo
	33	124.5	5.2	2061	21	AAA70228	Plasmodium falcipa
c	34	124	5.2	557	24	ABQ43458	Oligonucleotide fo
	35	124	5.2	557	24	ABQ43459	Oligonucleotide fo
c	36	124	5.2	5340	22	AAF26289	P. falciparum eg5
	37	123.5	5.2	6033	21	AAA70152	Plasmodium falcipa
	38	122.5	5.2	4311	21	AAA70133	Plasmodium falcipa
	39	122	5.1	5145	21	AAA70209	Plasmodium falcipa
	40	121.5	5.1	412	21	ABN81085	Shrimp polynucleot
	41	120.5	5.1	198	17	AAT17204	DNA-spanner oligon
	42	120.5	5.1	198	17	AAT17205	DNA-spanner oligon
	43	120	5.0	7143	21	AAA70250	Plasmodium falcipa
c	44	119.5	5.0	567	21	AAA29550	HIV codon altered
	45	119	5.0	3282	24	ABN79865	Fungal ZBC gene se

ALIGNMENTS

RESULT 1
AAA63670
ID AAA63670 standard; cDNA; 1722 BP.
XX
AC AAA63670;
XX
DT 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
XX
KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; VRN2; ss.
XX
OS Arabidopsis sp.
XX
PN WO200046358-A2.
XX
PD 10-AUG-2000.

```
XX
PF   25-JAN-2000; 2000WO-GB00197.
XX
PR   05-FEB-1999;   99GB-0002660.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Johanson U,  West J,  Dean C;
XX
DR   WPI; 2000-532899/48.
XX
PT   New nucleic acid derived from the FRI locus of a plant, e.g.
PT   Arabidopsis, encoding a polypeptide capable of specifically altering
PT   the flowering time of a plant -
XX
PS   Example 5; Page 53-54; 73pp; English.
XX
CC   The present sequence represents a VRN2 sequence from variety Landsberg
CC   erecta. The specification describes a sequence which encodes a
CC   polypeptide capable of specifically altering the flowering time of a
CC   plant. The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
CC   of Arabidopsis. The FRI polynucleotide is used to transform plants, so
CC   that the flowering time of a plant is altered. This is used, for example,
CC   for plants in which the leaves or tubers are a commercial product, where
CC   it is desirable to avoid 'blotting' (initiation of flowers and stem
CC   elongation) at too early a stage. Conversely, it may be desirable to
CC   alter flowering under certain circumstances e.g. to vary flower
CC   production across the seasons.
XX
SQ   Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

Alignment Scores:
Pred. No.:                3.53e-243      Length:          1722
Score:                    2378.00        Matches:         445
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       21             Gaps:            0

US-09-890-220-2 (1-445) x AAA63670 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 2
AAA47751
ID   AAA47751 standard; cDNA; 1722 BP.
XX
AC   AAA47751;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
```

```
XX
OS   Arabidopsis thaliana var 'Landsberg erecta'.
XX
FH   Key             Location/Qualifiers
FT   CDS             231..1568
FT                   /*tag=  a
FT                   /product= "VRN2 polypeptide"
FT   mutation        replace(1199,A)
FT                   /*tag=  b
FT                   /note= "found in mutant vin2-1; alters a Trp codon
FT                          to a STOP codon leading to a truncation of
FT                          the protein"
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00060.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 4; Fig 6; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

Alignment Scores:
Pred. No.:              3.53e-243       Length:         1722
Score:                  2378.00         Matches:        445
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     21              Gaps:           0

US-09-890-220-2 (1-445) x AAA47751 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565
```

```
RESULT 3
AAA63671
ID   AAA63671 standard; cDNA; 1715 BP.
XX
AC   AAA63671;
XX
DT   04-DEC-2000  (first entry)
XX
DE   Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
XX
KW   H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW   flower initiation; stem elongation; flower production; VRN2; ss.
XX
OS   Arabidopsis sp.
XX
PN   WO200046358-A2.
XX
PD   10-AUG-2000.
XX
PF   25-JAN-2000; 2000WO-GB00197.
XX
PR   05-FEB-1999;   99GB-0002660.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Johanson U,  West J,  Dean C;
XX
DR   WPI; 2000-532899/48.
XX
PT   New nucleic acid derived from the FRI locus of a plant, e.g.
PT   Arabidopsis, encoding a polypeptide capable of specifically altering
PT   the flowering time of a plant -
XX
PS   Example 5; Page 54-55; 73pp; English.
XX
CC   The present sequence represents a VRN2 sequence from variety Columbia.
CC   The specification describes a sequence which encodes a polypeptide
CC   capable of specifically altering the flowering time of a plant.
CC   The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
CC   of Arabidopsis. The FRI polynucleotide is used to transform plants, so
CC   that the flowering time of a plant is altered. This is used, for example,
CC   for plants in which the leaves or tubers are a commercial product, where
CC   it is desirable to avoid 'blotting' (initiation of flowers and stem
CC   elongation) at too early a stage. Conversely, it may be desirable to
CC   alter flowering under certain circumstances e.g. to vary flower
CC   production across the seasons.
XX
SQ   Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;

Alignment Scores:
Pred. No.:              2.88e-233       Length:         1715
Score:                  2285.00         Matches:        428
Percent Similarity:     97.51%          Conservative:   3
Best Local Similarity:  96.83%          Mismatches:     9
Query Match:            96.09%          Indels:         2
DB:                     21              Gaps:           1

US-09-890-220-2 (1-445) x AAA63671 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||  ||||||||||||||||||||  ||||||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||  ||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||:::|||         |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546
```

```
Qy       441 AspLys 442
             ||||||
Db      1547 GACAAG 1552

RESULT 4
AAA47753
ID   AAA47753 standard; cDNA; 1715 BP.
XX
AC   AAA47753;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana var 'Columbia'.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00061.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 5; Page 71; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;

Alignment Scores:
Pred. No.:              2.88e-233       Length:         1715
Score:                  2285.00         Matches:        428
Percent Similarity:     97.51%          Conservative:   3
Best Local Similarity:  96.83%          Mismatches:     9
Query Match:            96.09%          Indels:         2
DB:                     21              Gaps:           1

US-09-890-220-2 (1-445) x AAA47753 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              |||||||||||   ||||||||||||||||||||   ||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||   ||||||||||||||||||||||||||||||   ||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||:::|||      |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486
```

```
Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLys 442
              ||||||
Db       1547 GACAAG 1552

RESULT 5
AAA47755
ID   AAA47755 standard; cDNA; 1737 BP.
XX
AC   AAA47755;
XX
DT   16-NOV-2000  (first entry)
XX
DE   Aberrantly spliced VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana var 'Columbia'.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00062.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 8; Page 74-75; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant. This cDNA was generated from an aberrantly spliced VRN2
CC   gene.
XX
SQ   Sequence 1737 BP; 543 A; 330 C; 359 G; 505 T; 0 other;

Alignment Scores:
Pred. No.:                1.96e-230       Length:          1737
Score:                    2258.50         Matches:         428
Percent Similarity:       95.78%          Conservative:    3
Best Local Similarity:    95.11%          Mismatches:      9
Query Match:              94.97%          Indels:          10
DB:                       21              Gaps:            2

US-09-890-220-2 (1-445) x AAA47755 (1-1737)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 Lys----------------------GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
              |||                      ||||||||||||||||||||||||||||||||||||
Db        533 AAGGTGGGCAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTT 592

Qy        113 uPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        593 ATTTGAATTTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACT 652

Qy        133 uAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSerLe 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCT 712

Qy        153 uCysSerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLy 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAA 772

Qy        173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh 193
              |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        773 AGTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGC 832

Qy        193 rLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGl 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 CCTGCTGAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACA 892

Qy        213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATTTGAGATGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGC 952

Qy        233 aLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAl 253
              ||||||||||||||::|||||||||||||||||||||||||||||||||||||||||||
Db        953 TAAAGTTATATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGC 1012

Qy        253 aGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHi 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCA 1072

Qy        273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVa 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 CAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGT 1132

Qy        293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAA 1192

Qy        313 nLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIl 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTAT 1252

Qy        333 eAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLe 353
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db       1253 AGCAGATGGTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTT 1312

Qy        353 uHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHi 373
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 GCACTGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCA 1372

Qy        373 sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAA 1432
```

```
Qy        393 nSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAs 413
              |:::|||         |||   ||||||||||||||||||||   ||||||||||||||
Db       1433 TACCTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAA 1486

Qy        413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs 433
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1487 CACCAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGA 1546

Qy        433 pAsnLysAspAsnAsnSerArgAspLys 442
              |||||||||||||||||||||||||||
Db       1547 CAACAAGGACAATAACAGCAGAGACAAG 1574
```

RESULT 6
AAC51581
ID AAC51581 standard; DNA; 1497 BP.
XX
AC AAC51581;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69038.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	9.16e-195	Length:	1497
Score:	1923.00	Matches:	370
Percent Similarity:	90.10%	Conservative:	3

Best Local Similarity:	89.37%	Mismatches:	7
Query Match:	80.87%	Indels:	34
DB:	21	Gaps:	3

US-09-890-220-2 (1-445) x AAC51581 (1-1497)

```
Qy         43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 CCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCA 381

Qy         63 ArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThr 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 AGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACT 441

Qy         83 GluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGly 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGG 501

Qy        103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG--------- 552

Qy        123 GluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGluGly 142
                 |||   :::|||||||   ::::::        |||
Db        553 ---TACATCTCTGTAAATCCAACCATAATGTTTTGCAGC---------------------- 588

Qy        143 SerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgArgGln 162
                                                  ||||||||||||||||||||||||
Db        589 ----------------------------------TCGAAACCTCGTAAGCGGAGACAA 612

Qy        163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCC 672

Qy        183 SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeu 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 AGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTA 732

Qy        203 GlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPro 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 GGATATCCCGAGGCAACAGAACTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCA 792

Qy        223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVal 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTG 852

Qy        243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg--------- 259
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 GTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGTACTGCC 912

Qy        260 ------------------------SerHisLeuLeuLeuGlnLysArgGlnPheTyrHis 271
                                       ||||||||||||||||||||||||||||||||||||
Db        913 ATATTAACCGAGAAAATTTCTTCCAGCCACCTACTTCTTCAGAAACGCCAATTCTATCAT 972

Qy        272 SerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAsp 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 TCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGAT 1032

Qy        292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAsp 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 GAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGAT 1092

Qy        312 ValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArg 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 GTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGG 1152

Qy        332 ValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLys 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 GTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAA 1212

Qy        352 GluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrp 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1272
```

```
Qy        372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGG 1332

Qy        392 ArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAsp 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 CGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGAC 1392

Qy        412 SerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsn 431
              |||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db       1393 TCAAACACCAACAACAATAACATTGTGAGTCATCCCAATGACATAAACAACAAGAACAAT 1452

Qy        432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
              |||||||||||||||||||||||||||||||||||||||||||
Db       1453 GTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAA 1494
```

```
RESULT 7
AAA47752
ID   AAA47752 standard; DNA; 6338 BP.
XX
AC   AAA47752;
XX
DT   16-NOV-2000  (first entry)
XX
DE   VRN2 nucleic acid.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ds.
XX
OS   Arabidopsis thaliana var 'Landsberg erecta'.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Claim 9; Page 68-70; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant.
XX
SQ   Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;
```

```
Alignment Scores:
Pred. No.:              3.76e-182     Length:          6338
Score:                  1813.50       Matches:         440
Percent Similarity:     43.58%        Conservative:    1
Best Local Similarity:  43.48%        Mismatches:      4
Query Match:            76.26%        Indels:          569
DB:                     21            Gaps:            13

US-09-890-220-2 (1-445) x AAA47752 (1-6338)
```

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 2401

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2461

Qy         41 GlyAsn---------------------------------------------------------- 42
              ||||||
Db       2462 GGCAACGTATGATTTGCCTTCCTCTCTCATCATATTAGCTCAGTAATCTTTCATCTCCTG 2521

Qy         42 ---------------------------------------------------------------- 42

Db       2522 TGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGATTCATGGCGA 2581

Qy         43 -----------------------------------------------------------ProSerPhe 45
                                                                         |||||||||
Db       2582 GTGTGTGCTTCTTTTGTCTCCTAAATTTGAACTTGTTGTTTGTTGTTGCAGCCATCGTTT 2641

Qy         46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--------- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 CTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTA-TGTTTTCTT 2700

Qy         62 ---------------------------------------------------------------- 62

Db       2701 CTTGAATGTAGCTGCTACAGTGATATGTTATTTATCTTACTTCTAATATGGAAGCTGATG 2760

Qy         62 ---------------------------------------------------------------- 62

Db       2761 ACCTATTTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTGTTCATG 2820

Qy         63 -------------------------------------------ArgSerThrGlyMetVal 68
                                                         ||||||||||||||||||
Db       2821 CTATAAACTTACATTTTATAAAATTGTGTTTTGGTTAGGTCAAGATCTACTGGGATGGTA 2880

Qy         69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------------- 83
              |||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGGTTAGTCTTTTTCT 2940

Qy         83 ---------------------------------------------------------------- 83

Db       2941 GTTCGTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTATTTACTA 3000

Qy         84 ----------------------ValArgGluAspCysSerCysProPheCysSerMetLe 96
                                    ||||||||||||||||||||||||||||||||||||
Db       3001 TTTTTTTCTGTATTGTCATGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCT 3060

Qy         96 uCysGlySerPheLys------------------------------------------- 101
              ||||||||||||||||
Db       3061 ATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCTTTCATA 3120

Qy        101 ---------------------------------------------------------------- 101

Db       3121 TCTAACACTGTGAAATGCTACTGCTGTTTCATGCTGTATACTTTCACTGTTTGGTTACAT 3180

Qy        101 ---------------------------------------------------------------- 101

Db       3181 ATTTTTGTGTTTGTTGTTTGTCTTCTCACTCTTTTCGAACTGCTGAGTGTGTGCTTATCT 3240

Qy        102 -----------------------------------------------------------GlyLeuG 104
                                                                          |||||||
Db       3241 GAGAAAACATGTCCCAGATGGAGCTTACAACCAATTGTCTTGTGTCTATGCAGGGGCTGC 3300

Qy        104 lnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu---------- 120
              ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       3301 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATGTGGTTTTA 3360

Qy        120 ---------------------------------------------------------------- 120

Db       3361 TGGAAATTCTTGATTTGCTATGCCTTTATTAATGAGGTTATAGTTAAAAAAGGGTCTTTC 3420

Qy        121 -----------Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
```

```
Db       3421 CTATTGTAGCTTTCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTC 3480

Qy        137 IlePheGlu---------------------------------------------------- 139

Db       3481 ATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCGATAGCTGGTGAAAAT 3540

Qy        140 --------------------------------GluGluGlySerAspAspAspLysPheGl 149

Db       3541 TTTGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGACGATAAATTTGA 3600

Qy        149 uProPheSerLeu------------------------------------------------- 153

Db       3601 GCCCTTCTCTCTCTGGTAACCCTCAGAACCCCTTCGATTAAATACCTTAATAGCAGTAAC 3660

Qy        154 ----------------------------------------------------CysSerLy 156

Db       3661 TCCTTGCTTCTCTTGTCAGTACATCTCTGTAAATCCAACCATAATGTTTTGCAGCTCGAA 3720

Qy        156 sProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPh 176

Db       3721 ACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTT 3780

Qy        176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAs 196

Db       3781 TTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAA 3840

Qy        196 nAsp----------------------------------------------------------- 197

Db       3841 TGATGGTAAAATCATATCTTCTTCTGTGCGTTCCTTGTGGCTTAGAACTTCATATTACAG 3900

Qy        197 ----------------------------------------------------------- 197

Db       3901 AAGAAGATACAATGGCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTCTTCTTGCGA 3960

Qy        198 ----------------------------------------------GlyAsnArgGlyLe 202

Db       3961 GGGTATTGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAACCGTGGTTT 4020

Qy        202 uGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPr 222

Db       4021 AGGATATCCCGAGGCAACAGARCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACC 4080

Qy        222 oAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVa 242

Db       4081 AGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGT 4140

Qy        242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg-------- 259

Db       4141 GGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGGTTTGT 4200

Qy        259 ----------------------------------------------------------- 259

Db       4201 TCATCATGACACCCCGTCATCATAATTACCATTCCTGTTGTTACAAATGTTCTTCCTATT 4260

Qy        260 ----------------------------------------------SerHisLeuLe 263

Db       4261 ATGGATAAGTGTTTATAGTACTGCCATATTAACCGAGAAAATTTCTTCCAGCCACCTACT 4320

Qy        263 uLeuGlnLysArgGlnPheTyrHisSerHisArgVal------------------------ 275

Db       4321 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGTTCCTTCACC 4380

Qy        275 ----------------------------------------------------------- 275

Db       4381 TACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGCATACTTGGT 4440

Qy        275 ----------------------------------------------------------- 275

Db       4441 TTGTTCTCATCCAAATTTGTATTTTGATCTGTATGTATCAACGCAAAATAGTTATGTCCA 4500

Qy        276 ----------------------------------------------Gln-ProMetA 279

Db       4501 TGTTGTCTCCGTTTTATTGCCACTAACCAAAAAATGCATGTTTCTGTGACAAGCCAATGG 4560

Qy        279 laLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaA 299

Db       4561 CGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAG 4620

Qy        299 spPheGluAspArg----------------------------------------------- 303

Db       4621 ATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAATAGACAACAGA 4680

Qy        304 ----------------------------------------------------GlnMetLeu 306

Db       4681 AAATGGTATATGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGCTGCAGATGCTT 4740

Qy        307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326

Db       4741 GATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTT 4800

Qy        327 ValArgLysGlnArgVal-------------------------------------------- 332

Db       4801 GTAAGAAAACAAAGGTA-ACTACTTCTCTTACACATGAACAGACACAAAAAGACCTTATG 4859

Qy        332 ----------------------------------------------------------- 332

Db       4860 TCTTACATTCCATACCTGTCTAAATGATTTTGCTTATGGAACTTTGAGCTCAATTATGAT 4919

Qy        333 ------------------IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346

Db       4920 TGTTGATGTTTCAGGGTTATAGCAGATGGTCATATTTCTTGGGCATGTGAAGCATTTTCA 4979

Qy        347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp--------------- 361

Db       4980 AGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTAATATAAGTACACC 5039

Qy        361 ----------------------------------------------------------- 361

Db       5040 AAACATATACAGACACATAACTACACTATCAATCTTGTTTCGTTTTCTGAAAAAAAAATA 5099

Qy        362 ------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAsp 377

Db       5100 AAAATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGAC 5159

Qy        378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397

Db       5160 TCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACC 5219

Qy        398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417

Db       5220 ACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAAT 5279

Qy        418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437

Db       5280 AACATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAAT 5339

Qy        438 AsnSerArgAspLysValIleLys 445

Db       5340 AACAGCAGAGACAAAGTAATTAAA 5363
```

RESULT 8
AAA47754
ID AAA47754 standard; DNA; 5895 BP.
XX
AC AAA47754;
XX
DT 16-NOV-2000 (first entry)
XX
DE VRN2 nucleic acid.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time;
KW leaf size; leaf shape; shade avoidance response; reproduction;
KW breeding; pollination; cultivation; ds.
XX
OS Arabidopsis thaliana var 'Columbia'.

XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Claim 9; Page 72-74; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant.
XX
SQ Sequence 5895 BP; 1742 A; 1027 C; 1070 G; 2056 T; 0 other;

Alignment Scores:

Pred. No.:	9.98e-166	Length:	5895
Score:	1659.00	Matches:	424
Percent Similarity:	42.25%	Conservative:	4
Best Local Similarity:	41.86%	Mismatches:	12
Query Match:	69.76%	Indels:	577
DB:	21	Gaps:	14

US-09-890-220-2 (1-445) x AAA47754 (1-5895)

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1905 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 1964

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1965 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2024

Qy         41 GlyAsn---------------------------------------------------- 42
              |||
Db       2025 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 2083

Qy         42 ---------------------------------------------------------- 42

Db       2084 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 2143

Qy         43 ----------------------------------------------------------Pro 43
                                                                         |||
Db       2144 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 2203

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2204 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 2262

Qy         62 ---------------------------------------------------------- 62

Db       2263 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 2322

Qy         62 ---------------------------------------------------------- 62

Db       2323 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 2382

Qy         63 ----------------------------------------------ArgSerThrGly 66
                                                            ||||||||||||
Db       2383 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 2442

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              ||||||||||||||||||||||||||||||||||||||||:::||||||
Db       2443 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 2502

Qy         83 ---------------------------------------------------------- 83

Db       2503 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 2562

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           ||||||||||||||||||||||||||||||
Db       2563 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 2622

Qy         94 rMetLeuCysGlySerPheLys---------------------------------- 101
              ||||||||||||||||||||
Db       2623 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 2682

Qy        101 ---------------------------------------------------------- 101

Db       2683 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 2742

Qy        101 ---------------------------------------------------------- 101

Db       2743 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 2802

Qy        102 ----------------------------------------------------------G 102
                                                                         |
Db       2803 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 2862

Qy        102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       2863 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 2921

Qy        121 ---------------------------------------------------------- 121

Db       2922 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 2981

Qy        122 ----------------------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
                                     ||||||||||||||||||||||||||||||||||||
Db       2982 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 3041

Qy        134 snSerPheIlePheGlu------------------------------------------ 139
              ||||||||||||||||
Db       3042 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 3101

Qy        140 ----------------------------------------GluGluGlySerAspAspAsp 146
                                                      ||||||||||||||||||||
Db       3102 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 3161

Qy        147 LysPheGluProPheSerLeuCys---------------------------------- 154
              ||||||||||||||||||||
Db       3162 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 3220

Qy        154 ---------------------------------------------------------- 154

Db       3221 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 3280

Qy        155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3281 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 3340

Qy        174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db       3341 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 3400

Qy        194 LeuLeuAsnAsp------------------------------------------------ 197
              ||||||||||||
Db       3401 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 3460

Qy        197 ---------------------------------------------------------- 197
```

```
Db       3461 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 3520
Qy        198 ------------------------------------------------------GlyAs 199
                                                                       |||||
Db       3521 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 3580
Qy        199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3581 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 3640
Qy        219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db       3641 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 3700
Qy        239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3701 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 3760
Qy        259 g------------------------------------------------------------ 259
               |
Db       3761 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 3820
Qy        260 ------------------------------------------------------------Se 260
                                                                           ||
Db       3821 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 3880
Qy        260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----------- 276
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3881 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 3940
Qy        276 ------------------------------------------------------------ 276
Db       3941 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 4000
Qy        276 ------------------------------------------------------------ 276
Db       4001 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 4060
Qy        276 ------------------------------------------------------------ 276
Db       4061 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 4120
Qy        277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4121 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 4180
Qy        295 pAspValAlaAspPheGluAspArg------------------------------------ 303
               |||||||||||||||||||||||
Db       4181 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 4240
Qy        303 ------------------------------------------------------------ 303
Db       4241 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 4300
Qy        304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4301 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 4360
Qy        323 rpAsnSerPheValArgLysGln------------------------------------- 330
               |||||||||||||||||||||
Db       4361 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 4420
Qy        330 ------------------------------------------------------------ 330
Db       4421 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 4480
Qy        331 --------------------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
                                   ||||||||||||||||||||||||||||||||||||||
Db       4481 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 4540
Qy        344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
               ||||||||||||||||||||||||||||||||   ||||||||||||||||||
```

```
Db       4541 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 4600
Qy        361 ------------------------------------------------------------ 361
Db       4601 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 4660
Qy        362 ----------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
                                     ||||||||||||||||||||||||||||||||||||
Db       4661 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 4720
Qy        374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db       4721 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 4780
Qy        394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
               :|||         |||   ||||||||||||||||||||   ||||||||||||||||
Db       4781 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 4834
Qy        414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
               ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       4835 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 4894
Qy        434 nLysAspAsnAsnSerArgAspLys 442
               |||||||||||||||||||||||
Db       4895 CAAGGACAATAACAGCAGAGACAAG 4919
```

```
RESULT 9
AAZ36947
ID   AAZ36947 standard; cDNA; 2280 BP.
XX
AC   AAZ36947;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA encoding a MPC1 protein having flowering regulating activity.
XX
KW   Flowering regulating activity; MPC1; flowering; germination;
KW   super early flowering mutation; altered flowering time;
KW   flowering regulating gene; food crop; vegetable; flowering inhibition;
KW   productivity; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             310..2145
FT                   /*tag=  a
FT                   /product= "flowering regulating protein"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;   99EP-0305077.
XX
PR   26-JUN-1998;   98JP-0180065.
PR   24-JUN-1999;   99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC. .
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   P-PSDB; AAY53932.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 6; Page 16-21; 53pp; English.
XX
CC   The present sequence encodes a protein having a flowering regulating
CC   activity, which is designated MPC1. The genomic sequence is given
CC   in AAZ36948. A rice MPC1 is also disclosed in the specification. The
CC   rice and Arabidopsis cDNAs show significant homology with each other.
```

```
CC   A naturally occurring mutation of the MPC1 gene eliminates normal
CC   flowering regulating ability of plants, and leads to flowering
CC   immediately after germination (super early flowering mutation). The
CC   MPC1 polynucleotide sequence can be used to produce plants with altered
CC   flowering times in comparison with wild type plants, by enhancing
CC   or inhibiting the expression of the flowering regulating gene. Antisense
CC   polynucleotides can be used to reproduce the effects of the mutated MPC1
CC   gene. This alteration can be used to increase the yield of food crops.
CC   Flowering inhibition of vegetables increases their productivity.
XX
SQ   Sequence 2280 BP; 663 A; 445 C; 521 G; 651 T; 0 other;

Alignment Scores:
Pred. No.:                6.95e-88        Length:          2280
Score:                    923.00          Matches:         216
Percent Similarity:       43.23%          Conservative:    52
Best Local Similarity:    34.84%          Mismatches:      108
Query Match:              38.81%          Indels:          244
DB:                       21              Gaps:            9

US-09-890-220-2 (1-445) x AAZ36947 (1-2280)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||   ::::::   |||   :::   |||   ||||||   |||:::   :::|||
Db        310 ATGTGCCATGAAGACTCCCGTCTGCGTATTTCGGAAGAGGAGGAGATTGCTGCTGAAGAG 369

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||       ||||||||||||||   ||||||||||          |||::::::
Db        370 AGCTTGGCTGCCTATTGCAAGCCTGTTGAACTCTACAATATCATTCAACGCCGTGCTATT 429

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||   ||||||   ||||||||||:::|||||||||   ||||||   ||||||
Db        430 AGGAATCCCTTGTTTCTTCAGCGATGTTTGCATTATAAGATTGAGGCAAAACATAAAAGG 489

Qy         60 ------------------------------------------------------------ 60

Db        490 AGAATACAAATGACTGTATTCCTCTCGGGCGCTATAGATGCTGGGGTACAAACTCAAAAA 549

Qy         60 ------------------------------------------------------------ 60

Db        550 TTATTCCCTCTGTATATTTTGTTGGCAAGACTCGTTTCTCCTAAGCCTGTCGCTGAGTAT 609

Qy         60 ------------------------------------------------------------ 60

Db        610 TCTGCAGTATATAGGTTCAGTCGAGCATGTATCCTAACTGGTGGATTGGGGGTTGATGGA 669

Qy         60 ------------------------------------------------------------ 60

Db        670 GTTAGTCAAGCCCAAGCCAACTTTCTTCTCCCTGATATGAATAGACTCGCATTGGAGGCA 729

Qy         60 ------------------------------------------------------------ 60

Db        730 AAATCAGGATCACTCGCTATCTTGTTTATCAGCTTTGCTGGTGCGCAAAATTCTCAATTT 789

Qy         60 ------------------------------------------------------------ 60

Db        790 GGCATTGATTCAGGCAAGATTCATTCAGGAAATATAGGAGGACATTGTTTATGGAGCAAA 849

Qy         60 ------------------------------------------------------------ 60

Db        850 ATACCTCTGCAATCACTGTATGCGTCGTGGCAGAAATCACCAAACATGGACTTGGGACAG 909

Qy         61 ----------------------------------------------------LysSerArg 63
                                                                    ||||||
Db        910 AGAGTAGACACAGTCTCTCTTGTTGAAATGCAGCCTTGCTTCATAAAGCTAAAGTCCATG 969

Qy         64 Ser---------------------------------------------------------- 64
              |||
Db        970 AGTGAGGAAAAGTGTGTCTCGATTCAGGTGCCCAGCAATCCACTCACCTCGAGCTCTCCG 1029

Qy         64 ------------------------------------------------------------ 64

Db       1030 CAGCAAGTGCAAGTCACCATATCTGCAGAAGAAGTTGGGTCAACGGAAAAATCTCCTTAT 1089

Qy         64 ------------------------------------------------------------ 64

Db       1090 AGTTCATTTTCATATAATGACATCTCTTCCTCTTCCTTGTTGCAAATTATCAGGTTGAGA 1149

Qy         65 ThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGluVal 84
              ||||||   ||||||||||||||:::      ||||||   ||||||||||||||||||
Db       1150 ACAGGAAATGTAGTTTTCAACTACAGATACTATAACAACAAATTGCAGAAGACTGAAGTA 1209

Qy         85 ArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeuGln 104
              ||||||   ||||||||||||||   :::   |||   ||||||||||||||:::
Db       1210 ACTGAAGACTTTTCTTGTCCATTCTGCTTAGTAAAATGTGCCAGTTTCAAGGGCCTGAGA 1269

Qy        105 PheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGluTyr 124
              :::|||||   |||:::||||||||      ||||||||   :::   ||||||:::
Db       1270 TATCACTTGCCATCAACCCACGATCTCCTCAATTTCGAGTTTTGGGTAACTGAAGAATTT 1329

Qy        125 GlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGluGlySerAsp 144
              |||   ||||||||||||:::|||      :::   |||   :::          |||
Db       1330 CAGGCGGTAAATGTCTCCCTCAAGACTGAGACAATGATATCCAAGGTTAATGAGGATGAC 1389

Qy        145 ---AspAspLysPheGluProPheSerLeuCysSerLysPro---ArgLysArgArgGln 162
                 |||   |||   :::   |||         ||||||      |||:::||||||||
Db       1390 GTTGACCCAAAGCAGCAAACTTTCTTTTTTTCTTCCAAAAAATTCAGACGGAGGAGGCAA 1449

Qy        163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182
              :::          ::::::|||:::                          |||
Db       1450 AAGAGTCAGGTACGGAGCTCAAGGCAA-------------------------GGGCCT 1482

Qy        183 SerLeuThrAsnGly-----------------ThrGluAsnGlyIleThrLeuLeuAsn 196
              |||      |||                  |||::::::      ::::::   :::
Db       1483 CATCTTGGATTAGGTTGCGAGGTGCTAGATAAGACTGATGATGCTCATTCTGTTAGAAGT 1542

Qy        197 AspGlyAsnArg---------GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGln 213
              :::   :::|||         |||      |||            |||      |||
Db       1543 GAGAAGAGCCGAATACCACCTGGAAAGCATTAC-----------GAAAGAATTGGGGGT 1590

Qy        214 PheGluMetThrSerAsnIlePro-----------ProAlaIleAlaHisSerSerLeu 229
              |||           :::|||            ||||||        |||
Db       1591 GCTGAGTCTGGTCAAAGAGTTCCTCCTGGCACGAGTCCTGCAGACGTGCAATCATGTGGG 1650

Qy        230 AspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArg 249
              |||           :::           ::::::          |||||||||
Db       1651 GATCCAGATTATGTGCAGTCGATAGCTGGAAGTACAATGTTGCAGTTTGCAAAAACGAGG 1710

Qy        250 LysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPhe 269
              |||:::|||   ||||||||:::   |||:::      ||||||||||||||||||||
Db       1711 AAAATATCTATAGAACGGTCGGACTTGAGGAACCGAAGCCTCCTTCAGAAGAGACAGTTC 1770

Qy        270 TyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSer 289
              :::|||||||||||||   |||||||||||||||||||||||:::|||||||||||||||
Db       1771 TTCCACTCTCATCGAGCTCAGCCCATGGCTCTAGAACAAGTACTTTCGGACCGGGATAGT 1830

Qy        290 GluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPhe 309
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       1831 GAAGATGAAGTTGATGATGATGTGGCAGATTTTGAAGATAGAAGGATGCTCGATGATTTC 1890

Qy        310 ValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLys 329
              |||||||||   ||||||||||||||   ||||||:::|||||||||||||||||||||
Db       1891 GTTGATGTGACTAAAGATGAGAAACAGATGATGCACATGTGGAACTCGTTTGTGAGGAAG 1950

Qy        330 GlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyr 349
              |||||||||:::|||||||||||||||   |||||||||||||||||||||||   :::
Db       1951 CAGCGAGTATTAGCAGATGGTCACATTCCATGGGCATGCGAGGCATTCTCAAGATTGCAC 2010

Qy        350 GluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLys 369
                    :::   |||         |||   ||||||||||||:::|||:::::|||
Db       2011 GGACCCATCATGGTTCGAACACCGCACTTGATTTGGTGCTGGAGAGTGTTTATGGTGAAA 2070

Qy        370 LeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
              ||||||||||||||||||:::|||:::   |||:::|||||||||||||||   ||||||
Db       2071 CTGTGGAACCACGGTCTTCTTGATGCCCGAACCATGAACAACTGTAATACCTTTCTCGAA 2130
```

```
RESULT 10
AAZ36953
ID   AAZ36953 standard; cDNA; 2248 BP.
XX
AC   AAZ36953;
XX
DT   13-MAR-2000  (first entry)
XX
DE   cDNA encoding an Os-MPC1 protein having flowering regulating activity.
XX
KW   Flowering regulating activity; MPC1; flowering; germination;
KW   super early flowering mutation; altered flowering time;
KW   flowering regulating gene; food crop; vegetable; flowering inhibition;
KW   productivity; ss.
XX
OS   Oryza sativa.
XX
FH   Key             Location/Qualifiers
FT   CDS             86..1900
FT                   /*tag= a
FT                   /product= "flowering regulating protein"
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;   99EP-0305077.
XX
PR   26-JUN-1998;   98JP-0180065.
PR   24-JUN-1999;   99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   P-PSDB; AAY53933.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Claim 6; Page 40-44; 53pp; English.
XX
CC   The present sequence encodes a protein having a flowering regulating
CC   activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
CC   disclosed in the specification. The rice and Arabidopsis cDNAs show
CC   significant homology with each other. A naturally occurring mutation
CC   of the MPC1 gene eliminates normal flowering regulating ability of
CC   plants, and leads to flowering immediately after germination (super
CC   early flowering mutation). The MPC1 polynucleotide sequence can be
CC   used to produce plants with altered flowering times in comparison
CC   with wild type plants, by enhancing or inhibiting the expression of
CC   the flowering regulating gene. Antisense polynucleotides can be used
CC   to reproduce the effects of the mutated MPC1 gene. This alteration
CC   can be used to increase the yield of food crops. Flowering inhibition
CC   of vegetables increases their productivity.
XX
SQ   Sequence 2248 BP; 641 A; 454 C; 493 G; 660 T; 0 other;

Alignment Scores:
Pred. No.:              3.91e-86      Length:          2248
Score:                  906.50        Matches:         209
Percent Similarity:     42.68%        Conservative:    59
Best Local Similarity:  33.28%        Mismatches:      109
Query Match:            38.12%        Indels:          251
DB:                     21            Gaps:            7

US-09-890-220-2 (1-445) x AAZ36953 (1-2248)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||        ||||||:::   |||||||:::|||   ::::::   :::|||
Db         86 ATGTGCCGCCACCAGCCAAGGGCTCGGCTCTCTCCCGATGAGCAGCTTGCAGCTGAAGAA 145

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::      :::|||||||||||||||   ||||||||||||         ||||||:::
Db        146 AGCTTCGCATTATACTGCAAGCCGGTCGAGTTGTATAATATCATTCAGCGCCGATCCATT 205

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||:::||||||   |||||||||   |||||||||   |||:::|||||||:::
Db        206 AAAAATCCTGCTTTTCTTCAAAGATGCCTTCTTTACAAGATTCACGCAAGACGGAAGAAG 265

Qy         61 LysSerArgSerThr---------------------------------------------- 65
              :::|||      |||
Db        266 AGGAGCCTGATAACCATATCACTTTCTGGAGGCACAAATAAAGAACTGCGGGCACAAAAT 325

Qy         65 ---------------------------------------------------------------- 65

Db        326 ATCTTTCCTCTTTATGTTCTGTTAGCTAGACCTACTAATAATGTTTCACTTGAAGGGCAT 385

Qy         65 ---------------------------------------------------------------- 65

Db        386 TCTCCGATATATCGATTCAGTCGTGCTTGTTTGTTGACTTCTTTTCATGAATTTGGAAAT 445

Qy         65 ---------------------------------------------------------------- 65

Db        446 AAAGACTACACTGAAGCAACATTCGTCATTCCTGATGTGAAGAACTTAGCAACCTCCCGA 505

Qy         65 ---------------------------------------------------------------- 65

Db        506 GCTTGCAGCCTTAATATTATCCTTATCAGCTGTGGACGAGCTGAGCAAACTTTTGATGAC 565

Qy         65 ---------------------------------------------------------------- 65

Db        566 AATAACTGTTCTGGGAACCATGTGGAAGGCTCTACTCTCCAAAAGCTTGAAGGGAAGTGT 625

Qy         65 ---------------------------------------------------------------- 65

Db        626 TTCTGGGGTAAAATACCAATCGATCTTCTTGCTTCATCTTTGGGAAATTGTGTGAGCTTA 685

Qy         65 ---------------------------------------------------------------- 65

Db        686 AGTTTGGGACATACCGTGGAAATGTCTTCCACGGTTGAGATGACCCCAAGCTTCTTAGAG 745

Qy         65 ---------------------------------------------------------------- 65

Db        746 CCAAAATTTCTGGAGGATGACAGTTGCTTGACATTTTGCTCTCAGAAGGTTGATGCTACT 805

Qy         65 ---------------------------------------------------------------- 65

Db        806 GGTTCATTTCAACTGCAAGTTAGCATATCTGCTCAAGAGGCTGGTGCAAAAGACATGTCC 865

Qy         65 ---------------------------------------------------------------- 65

Db        866 GAGTCTCCTTATAGTGTTTATTCATATAATGATGTGCCACCTTCGTCATTGACACATATT 925

Qy         66 ----------GlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
                        |||   |||:::||||||||||||      ||||||||||:::|||
Db        926 ATAAGGTTGAGATCTGGCAATGTGCTTTTTAACTACAAATACTACAATAATACTATGCAA 985

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||||   :::   ||||||||||||
Db        986 AAAACCGAAGTCACTGAAGATTTTTCTTGCCCATTTTGCTTGGTACCATGTGGCAGCTTT 1045

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||       |||||||||:::|||||||||||||||   :::||||||   :::
Db       1046 AAGGGTCTAGGATGTCACCTAAACGCATCGCATGACCTTTTCCATTATGAGTTTTGGATA 1105

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu--- 139
              ||||||   |||   ||||||||||||:::|||   :::|||:::      |||
Db       1106 TCTGAAGAGTGCCAGGCTGTTAATGTTAGTCTGAAGACTGATTCTTGGAGAACAGAGCTT 1165

Qy        140 ---GluGluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArg 158
                 ||||||      |||   :::   :::   ||||||     |||:::   :::
Db       1166 TTGGCTGAGGGAGTT---GATCCAAGACATCAAACATTTTCGTACCGCTCAAGATTTAAG 1222

Qy        159 LysArg---ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
```

```
Db       1223 AAGCGTAAAAGGGTGGAAATCTCAAGTGATAAAATTAGGCATGTACATCCACATATTGTG 1282

Qy        178 ProLeuAspSerPro---SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsn 196

Db       1283 GATTCAGGATCACCTGAAGATGCCCAGGCAGGATCTGAAGACGATTACGTGCAGAGGGAA 1342

Qy        197 AspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 216

Db       1343 AATGGTAGT---------------------------------------------------- 1351

Qy        217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal--- 235

Db       1352 ----------------TCTGTAGCACACGCTTCTGTTGATCCTGCTAATTCATTACAC 1393

Qy        236 ------IleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAla 253

Db       1394 GGTAGCAATCTTTCAGCACCAACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTT 1453

Qy        254 GluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHis 273

Db       1454 GAACGAGCTGATCCCAGAAATCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCAC 1513

Qy        274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293

Db       1514 AGGGCTCAACCAATGGCATGGAGCAAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTT 1573

Qy        294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313

Db       1574 GATGATGACATTGCTGATTTTGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACA 1633

Qy        314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333

Db       1634 AAAGACGAGAAACTTATTATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTA 1693

Qy        334 AlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353

Db       1694 GCGGATGGCCATATTCCCTGGGCATGCGAAGCATTCTCGCAGTTTCATGGACAAGAACTT 1753

Qy        354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373

Db       1754 GTACAAAATCCAGCTCTACTATGGTGTTGGAGGTTTTTTATGGTCAAACTCTGGAACCAC 1813

Qy        374 GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393

Db       1814 AGTCTACTGGATGCGCGAGCCATGAATGCCTGCAACACAATTCTTGAAGGCTACCTGAAC 1873

Qy        394 SerSerAspThrThrThrThrAsn 401

Db       1874 GGAAGCTCGGATCCAAAGAAAAAT 1897

RESULT 11
AAA47758
ID   AAA47758 standard; cDNA; 558 BP.
XX
AC   AAA47758;
XX
DT   16-NOV-2000  (first entry)
XX
DE   At Hyp 2245035 (ATFCA7_4) (modified) cDNA.
XX
KW   Vernalization gene; VRN2; plant characteristic; flowering time;
KW   leaf size; leaf shape; shade avoidance response; reproduction;
KW   breeding; pollination; cultivation; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200044918-A1.
XX
PD   03-AUG-2000.
XX
PF   28-JAN-2000; 2000WO-GB00248.
XX
PR   28-JAN-1999;   99GB-0001927.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dean C,  Gendall A;
XX
DR   WPI; 2000-499333/44.
DR   P-PSDB; AAB00065.
XX
PT   Isolated vernalization gene VRN2 is used to produce transgenic plants
PT   with altered vernalization response, flowering time, leaf size and/or
PT   shape or shade avoidance response for maximized reproductive success
XX
PS   Disclosure; Page 76; 105pp; English.
XX
CC   Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC   plant encode polypeptides which are capable of affecting one or more
CC   vernalization responses such as, flowering time, leaf size and/or
CC   shape or the shade avoidance response of a plant into which the
CC   nucleic acid is introduced. Introducing such sequences into plants
CC   to alter these characteristics maximises the reproductive success of
CC   the plant. This cDNA represents another Arabidopsis thaliana sequence
CC   having homology to the VRN2 gene.
XX
SQ   Sequence 558 BP; 191 A; 122 C; 112 G; 133 T; 0 other;

Alignment Scores:
Pred. No.:              1.55e-38        Length:         558
Score:                  451.50          Matches:        99
Percent Similarity:     67.18%          Conservative:   32
Best Local Similarity:  50.77%          Mismatches:     43
Query Match:            18.99%          Indels:         21
DB:                     21              Gaps:           7

US-09-890-220-2 (1-445) x AAA47758 (1-558)

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256

Db         16 CTGACAACAGAAGCTAAGGTCCCT----------------------GCTAAGCGATCA 51

Qy        257 GluAlaArgSerHisLeuLeu---LeuGlnLysArgGlnPheTyrHisSerHisArgVal 275

Db         52 AAGGCTACAAGCCACTACTTGCCTCTTCATAAACGCCAGTTCTATCATTCCCGAACCGGT 111

Qy        276 GlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAsp--- 294

Db        112 CAGCCATTGTCACTTGAGCAAGTTATGTCTGACCGAGATAGCGAAAATGACGTCGACAAA 171

Qy        295 ---AspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313

Db        172 AATGATGATGCTGCACATCTCGAAGAAAGCCAGATGCTTAATGGTTCCATGGATGAGAAT 231

Qy        314 Lys---AspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgVal 332

Db        232 GAAATCGTAGCAGAGAGATTCATAAAACTTTGGAACTCCTTTGTTAAACAGCAAAGGATT 291

Qy        333 IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGlu 352

Db        292 GTTGCAGATGCTCATATTCCTTGGGCATGTGAAGCATTCTCAAGATTACACCTGCAAGAG 351

Qy        353 LeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372

Db        352 CTGCGCAGTAACTTATCACTCGACTTGTGCTGGAGACAATTCATGATCAAACAATGGGAT 411

Qy        373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392

Db        412 TATGGACTTCTTGACAGAGTCACCATGAACAAATGCAATACCATCATCTACCAT------ 465

Qy        393 AsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSer 412

Db        466 AATATCTCAACTACCAACGATGACATAAACAATAACAACACA--------------AGG 510

Qy        413 AsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsn 427
```

```
                    |||:::|||   ::::::|||||||      :::|||||||||
Db          511 ACGACTGATAATATGGATGTTGTCGAC------GATGACATAAAC 549

RESULT 12
AAZ36948
ID   AAZ36948 standard; DNA; 5580 BP.
XX
AC   AAZ36948;
XX
DT   13-MAR-2000  (first entry)
XX
DE   DNA encoding a MPC1 protein having flowering regulating activity.
XX
KW   Flowering regulating activity; MPC1; flowering; germination;
KW   super early flowering mutation; altered flowering time;
KW   flowering regulating gene; food crop; vegetable; flowering inhibition;
KW   productivity; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             588..5545
FT                   /*tag= a
FT                   /product= "flowering regulating protein"
FT                   /note= "contains introns"
FT   exon            588..713
FT                   /*tag= b
FT                   /number= 1
FT   intron          714..930
FT                   /*tag= c
FT                   /number= 1
FT   exon            931..986
FT                   /*tag= d
FT                   /number= 2
FT   intron          987..1132
FT                   /*tag= e
FT                   /number= 2
FT   exon            1133..1247
FT                   /*tag= f
FT                   /number= 3
FT   intron          1248..1344
FT                   /*tag= g
FT                   /number= 3
FT   exon            1345..1504
FT                   /*tag= h
FT                   /number= 4
FT   intron          1505..1596
FT                   /*tag= i
FT                   /number= 4
FT   exon            1597..1647
FT                   /*tag= j
FT                   /number= 5
FT   intron          1648..1854
FT                   /*tag= k
FT                   /number= 5
FT   exon            1855..1994
FT                   /*tag= l
FT                   /number= 6
FT   intron          1995..2197
FT                   /*tag= m
FT                   /number= 6
FT   exon            2198..2260
FT                   /*tag= n
FT                   /number= 7
FT   intron          2261..2350
FT                   /*tag= o
FT                   /number= 7
FT   exon            2351..2472
FT                   /*tag= p
FT                   /number= 8
FT   intron          2473..2714
FT                   /*tag= q
FT                   /number= 8
FT   exon            2715..2779
FT                   /*tag= r
FT                   /number= 9
FT   intron          2780..2870
FT                   /*tag= s
FT                   /number= 9
FT   exon            2871..2930
FT                   /*tag= t
FT                   /number= 10
FT   intron          2931..3038
FT                   /*tag= u
FT                   /number= 10
FT   exon            3039..3092
FT                   /*tag= v
FT                   /number= 11
FT   intron          3093..3174
FT                   /*tag= w
FT                   /number= 11
FT   exon            3175..3234
FT                   /*tag= x
FT                   /number= 12
FT   intron          3235..3654
FT                   /*tag= y
FT                   /number= 12
FT   exon            3656..3701
FT                   /*tag= z
FT                   /number= 13
FT   intron          3702..3784
FT                   /*tag= aa
FT                   /number= 13
FT   exon            3785..3885
FT                   /*tag= ab
FT                   /number= 14
FT   intron          3886..4052
FT                   /*tag= ac
FT                   /number= 14
FT   exon            4053..4272
FT                   /*tag= ad
FT                   /number= 15
FT   intron          4273..4428
FT                   /*tag= ae
FT                   /number= 15
FT   exon            4429..4477
FT                   /*tag= af
FT                   /number= 16
FT   intron          4478..4552
FT                   /*tag= ag
FT                   /number= 16
FT   exon            4553..4636
FT                   /*tag= ah
FT                   /number= 17
FT   intron          4637..4982
FT                   /*tag= ai
FT                   /number= 17
FT   exon            4983..5062
FT                   /*tag= aj
FT                   /number= 18
FT   mutation        replace (5039, A)
FT                   /*tag= ak
FT                   /note= "this mutation destroys the flowering
FT                        regulating ability of the protein"
FT   intron          5063..5265
FT                   /*tag= al
FT                   /number= 18
FT   exon            5266..5355
FT                   /*tag= am
FT                   /number= 19
FT   intron          5356..5445
FT                   /*tag= an
FT                   /number= 19
FT   exon            5446..5542
FT                   /*tag= ao
FT                   /number= 20
```

```
XX
PN   EP967278-A2.
XX
PD   29-DEC-1999.
XX
PF   28-JUN-1999;   99EP-0305077.
XX
PR   26-JUN-1998;   98JP-0180065.
PR   24-JUN-1999;   99JP-0179043.
XX
PA   (MITA ) MITSUI CHEM INC.
XX
PI   Yoshida N,  Kato Y,  Takahashi S,  Yanai Y,  Hiratsuka J,  Miwa T;
XX
DR   WPI; 2000-064612/06.
DR   P-PSDB; AAY53932.
XX
PT   Novel DNA used to produce transgenic plants with altered floral
PT   regulation which can have increased crop yields -
XX
PS   Example 1; Page 26-34; 53pp; English.
XX
CC   The present sequence encodes a protein having a flowering regulating
CC   activity, which is designated MPC1. The cDNA sequence is given
CC   in AAZ36947. A rice MPC1 is also disclosed in the specification. The
CC   rice and Arabidopsis cDNAs show significant homology with each other.
CC   A naturally occurring mutation of the MPC1 gene eliminates normal
CC   flowering regulating ability of plants, and leads to flowering
CC   immediately after germination (super early flowering mutation). The
CC   MPC1 polynucleotide sequence can be used to produce plants with altered
CC   flowering times in comparison with wild type plants, by enhancing
CC   or inhibiting the expression of the flowering regulating gene. Antisense
CC   polynucleotides can be used to reproduce the effects of the mutated MPC1
CC   gene. This alteration can be used to increase the yield of food crops.
CC   Flowering inhibition of vegetables increases their productivity.
XX
SQ   Sequence 5580 BP; 1496 A; 999 C; 1075 G; 2010 T; 0 other;

Alignment Scores:
Pred. No.:              2.4e-32      Length:          5580
Score:                  407.50       Matches:         187
Percent Similarity:     24.79%       Conservative:    47
Best Local Similarity:  19.81%       Mismatches:      87
Query Match:            17.14%       Indels:          626
DB:                     21           Gaps:            18

US-09-890-220-2 (1-445) x AAZ36948 (1-5580)

Qy         65 ThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGluVal 84
              ||||||   ||||||||||||||||:::      ||||||   ||||||||||||||
Db       2722 ACAGGAAATGTAGTTTTCAACTACAGATACTATAACAACAAATTGCAGAAGACTGAAGG- 2780

Qy         84 ------------------------------------------------------------ 84

Db       2781 TAACTAGTATTATTTTAACCTGTTTCATACCCATGTGTGTCTATATTTCATCCGTTACCC 2840

Qy         85 ---------------------------------------ArgGluAspCysSerCysPro 91
                                                     ||||||   |||||||||
Db       2841 TAACCTGTTACGTATATGTTTGCTATGTGTCTTGCAGTAACTGAAGACTTTTCTTGTCCA 2900

Qy         92 PheCysSerMetLeuCysGlySerPheLys---------------------------- 101
              ||||||   :::   |||   |||||||||
Db       2901 TTCTGCTTAGTAAAATGTGCCAGTTTCAAGGTGGACTTTCATTTCCATTCTCATTCATCC 2960

Qy        101 ------------------------------------------------------------ 101

Db       2961 TCTTAGTCAAAGATACAGCTGTAGTGACTAGTCTTTGTAGTGATGCAATCTTTTCTTTTT 3020

Qy        102 ------------------GlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlu 115
                               ||||||::::::||||||   |||:::|||||||||
Db       3021 CTCCCAATCATGTTGTAGGGCCTGAGATATCACTTGCCATCAACCCACGATCTCCTCAAT 3080

Qy        116 PheGlu-PheLysLeu---------------------------------------------- 120
```

```
                ||||||  |||   |||
Db       3081 TTCGAGTTTTGGGTTGTAGCTTTAAAATTCAGTTAACCTGTTTGATCTTTTTTTTTTATT 3140

Qy        121 ------------------------------------PheGluGluTyrGlnThrValAs 128
                                                  ||||||:::|||   |||||
Db       3141 TTGTGGGTGCCACTAATCTGCTTTACTTGGTTAGGTAACTGAAGAATTTCAGGCGGTAAA 3200

Qy        128 nValSerValLys---------------------------------------------- 132
               |||||||:::|||
Db       3201 TGTCTCCCTCAAGACTGAGACAATGATATCCAAGGTTAGAACATCTTGTTTGTTCGATTT 3260

Qy        132 ------------------------------------------------------------ 132

Db       3261 ATGTTCATTAGTTTCTCTGCTGTATATCTTATAGGCTGTAACAAATTCATTTTTCATTTA 3320

Qy        133 ------------------LeuAsnSerPheIlePheGluGluGluGlySer-------- 143
                               |||   :::|||::::::      :::|||:::
Db       3321 AACTAATATCCTCCATGGGTTGTTGACTTTTGTGTGGTTAAATAAGGGAACTGGAATCTT 3380

Qy        144 ------------------------------------------------AspAspAspLy 147
                                                               |||:::::::||
Db       3381 TAGTTGCTATTTGTCACACTATGATCCTTGCTATTGTCCTTAATAGCGTGATGAGAATAA 3440

Qy        147 sPheGlu---------------------------------------------------- 149
               |   :::
Db       3441 ACT-CAAAATGACATCGCTGTTCTGTTTACTTTTTGTGGCCATGAGACCGTCAAAGCTCG 3499

Qy        149 ------------------------------------------------------------ 149

Db       3500 ACTGTAGAATAAAGTCCTGGATTATATAGGAGTGTCAAATCTAATTGAAGTAGTTGGTTC 3559

Qy        149 ------------------------------------------------------------ 149

Db       3560 TACAATATATTCTATGTCTTTGTAGTTTTTCCTATTTGATGATTACTCTTAGCACAGTTT 3619

Qy        149 ------------------------------------------------------------ 149

Db       3620 TCTAAATGTTAATGTTCATTAAAAAATCTGCTCAGGTTAATGAGGATGACGTTGACCCAA 3679

Qy        150 ------------------------------------------------ProPhe-------- 151
                                                               ||||||
Db       3680 AGCAGCAAACTTTCTTTTTTTCGTAAGTTATCTGGCCTATATGTTGCCTTTTATTATCTT 3739

Qy        152 ------------------------------SerLeuCysSerLysProArgLysArgAr 161
                                           |||:::|||         :::|||   ||
Db       3740 TCCAGCATCTGTGTGAGACCATAAAAATTCTTCAATATGTGACAGTTCCAAAAAATTCAG 3799

Qy        161 gGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysVal-------------------- 174
               |:::|||      ::::::     |||      :::
Db       3800 ACGGAGGAGGCAAAAGAGTCAGGTACGGAGCTCAAGGCAAGGGCCTCATCTTGGATTAGG 3859

Qy        175 -CysPheLeu------ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIl 192
                |||   :::           |||         ||||||      |||   |||   ||
Db       3860 TTGCGAGGTGCTAGATAAGACTGATGGTATGTGTTTGACTGAAATGACAGTTAATTGGAT 3919

Qy        192 eThrLeuLeu------------------------------------------------- 195
               |      :::
Db       3920 TTGTAGTATTGGCTTCTTTTGTGATGAGAGCCTGTCTTAGTTGTATATTTTACGAGTATT 3979

Qy        196 ------------------------------AsnAspGlyAsnArgGlyLeuGlyTyrPr 205
                                           |||   |||:::   |||      :::
Db       3980 TTACTTTGTTATGTGCAATTTTGCATGCAACAACGTTGGATCATTTGGCACAGCTTTTTA 4039

Qy        205 oGluAlaThrGluLeuAlaGlyGlnPheGluMet--ThrSerAsnIleProProAlaIle 224
                      ::::::           |||:::     |||   |||||||||
Db       4040 TTCTTACTTTCAGATGCTCATTCTGTTAGAAGTGAGAAGAGCCGAATACCACCTGGA--- 4096

Qy        225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValPro 244
               |||               ||||||:::         :::   :::   |||   |||
Db       4097 AAGCATTACGAAAGAATTGGGGGTGCTGAG---------TCTGGTCAAAGAGTTCCTCCT 4147

Qy        245 AlaThr---------------------------------------------------- 246
                  |||
```

```
Db       4148 GGCACGAGTCCTGCAGACGTGCAATCATGTGGGGATCCAGATTATGTGCAGTCGATAGCT 4207
Qy        247 ------------------------LysThrArgLysLeuSerAlaGluArgSerGluAla 258
                                      ||||||||||||:::|||   |||||||||:::
Db       4208 GGAAGTACAATGTTGCAGTTTGCAAAAACGAGGAAAATATCTATAGAACGGTCGGACTTG 4267

Qy        259 ArgSerHisLeu---------------------------------------------- 262
               |||   ::::::
Db       4268 AGGAAGTATGTTTGACTTCCTTTTGTCGTTCTATCCTCTTCTTCAATTTATATTTAACTA 4327

Qy        262 ------------------------------------------------------------ 262

Db       4328 CATATGGTTCATGCATGAAAAATTGTGTCCTAGTTTTATAACAAGTAGCTTGTTAATCCC 4387

Qy        263 -----------------------------------------------LeuLeuGlnLys 266
                                                            ||||||||||||
Db       4388 AAATGATGTGAGTGAGTTTTTCAAATTTTTTCCTCCTCCAGCCGAAGCCTCCTTCAGAAG 4447

Qy        267 ArgGlnPheTyrHisSerHisArgVal------------------------------- 275
               |||||||||:::||||||||||||
Db       4448 AGACAGTTCTTCCACTCTCATCGAGCTCAGGTGATCTTTTTTCTTTAGCTCTCTTGCTTT 4507

Qy        276 ---------------------------------------GlnProMetAlaLeuGlu 281
                                                    ||||||||||||||||
Db       4508 TGAAGATTGCAATTGATTTTGACTTTGCTATGTGTACTATCGCAGCCCATGGCTCTAGAA 4567

Qy        282 GlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGlu 301
               ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4568 CAAGTACTTTCGGACCGGGATAGTGAAGATGAAGTTGATGATGATGTGGCAGATTTTGAA 4627

Qy        302 AspArg---------------------------------------------------- 303
               ||||||
Db       4628 GATAGAAGGGTATGTTTTTGAATTTAATATTTTCACCGCATCAGTAGTTGGGTAGAATAA 4687

Qy        303 ------------------------------------------------------------ 303

Db       4688 AGCTCAGTAGTTGGGTAGATATATGTTTCATGTGAAAGGGAAAGGAATATTGAAGACTGG 4747

Qy        303 ------------------------------------------------------------ 303

Db       4748 GCATGGGCAAACGTTAGGAGCAATATTGTAATGGTTCAGAGATCAATAGAAAATATGTGA 4807

Qy        303 ------------------------------------------------------------ 303

Db       4808 GCAAGCCTCACGGTTTGATATGGAACAGTAGAACCAGATCATTAGTGCTTATATAACACT 4867

Qy        303 ------------------------------------------------------------ 303

Db       4868 CATTAAAAGACGAAGTGTGTCCGTTTGTACTCGATTCTAACATAGTTGATTCTAACATAG 4927

Qy        304 ---------------------------------------------------GlnMetLe 306
                                                                 ||||||||
Db       4928 TTTGTCTGATTCTCCATATAGTGAATAACGTTATTTCCTATTACTATTCTTTCAGATGCT 4987

Qy        306 uAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPh 326
               ||||||||||||||||||   ||||||||||||||   ||||||:::|||||||||||
Db       4988 CGATGATTTCGTTGATGTGACTAAAGATGAGAAACAGATGATGCACATGTGGAACTCGTT 5047

Qy        326 eValArgLysGlnArgVal-------------------------------------- 332
               |||||||||||||||
Db       5048 TGTGAGGAAGCAGCGGTA-TGTCTTATCTCTTTTCAGTACATGTCACGTGGAGTTTTCCA 5106

Qy        332 ------------------------------------------------------------ 332

Db       5107 GTATAAACATTTAGAGTCGCGCATGTAAAGGTTGTGGATAATTCCTGCCTGGGTTCTTCT 5166

Qy        332 ------------------------------------------------------------ 332

Db       5167 GGTTAAAAAAAAAAAACTGAACAATTAGATAACATACGCATCCATGTTCTCTGACTCATT 5226

Qy        333 -------------------------------------------IleAlaAspGlyHisIl 338
                                                        :::|||||||||||||||
Db       5227 ATAAGCATTACCTTGACAGTGGTTTTGGACCCTTTGCAGAGTATTAGCAGATGGTCACAT 5286

Qy        338 eSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSe 358
               |   ||||||||||||||||||||||||   :::   :::   |||
Db       5287 TCCATGGGCATGCGAGGCATTCTCAAGATTGCACGGACCCATCATGGTTCGAACACCGCA 5346

Qy        358 rLeuPheTrp------------------------------------------------- 361
                |||   |||
Db       5347 CTTGATTTGGTAATTCAACTCTCATTTCTTCCATTGTTTTTTCCAGTGTATCGGAGAAGA 5406

Qy        362 -----------------------------------------CysTrpArgLeuPheLeuIl 368
                                                      |||||||||:::|||:::::
Db       5407 AAGCGGTTTTGTTGATAAAAGTGAGCTTTTTTTGTGTAGGTGCTGGAGAGTGTTTATGGT 5466

Qy        368 eLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLe 388
               :|||||||||||||||||||:::|||:::   |||:::|||||||||||||||   ||
Db       5467 GAAACTGTGGAACCACGGTCTTCTTGATGCCCGAACCATGAACAACTGTAATACCTTTCT 5526

Qy        388 uGlu 389
               ||||
Db       5527 CGAA 5530
```

RESULT 13
AAA47757
ID AAA47757 standard; cDNA; 325 BP.
XX
AC AAA47757;
XX
DT 16-NOV-2000 (first entry)
XX
DE AI163743 Expressed sequence tag (modified) cDNA.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time; poplar;
KW leaf size; leaf shape; shade avoidance response; reproduction; EST;
KW breeding; pollination; cultivation; expressed sequecne tag; ss.
XX
OS Populus tremula.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
DR P-PSDB; AAB00064.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 76; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA represents expressed sequence tag clone AI163743
CC from poplar trees (Populus tremula x Populus tremuloides) which has
CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
XX
SQ Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;

Alignment Scores:
Pred. No.: 4.84e-30 Length: 325
Score: 368.50 Matches: 76

```
Percent Similarity:     68.00%          Conservative:  9
Best Local Similarity:  60.80%          Mismatches:    21
Query Match:            15.50%          Indels:        19
DB:                     21              Gaps:          2

US-09-890-220-2 (1-445) x AAA47757 (1-325)

Qy        204 TyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAla 223
              ||||||:::|||                |||            :::|||:::   |||
Db          8 TATCCTGATGCTGAA------TGTGCTCAATTGGTACCTGGGAATAATCTTGCACCT--- 58

Qy        224 IleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValVal 243
                                                                |||::::::
Db         59 ----------------------------------------------CCTGCCATGCTA 70

Qy        244 ProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeu 263
                  ||||||||||||||||||   |||||||||:::   |||:::      |||
Db         71 CAATTTGCAAAGACAAGAAAATTATCAATTGAACGGTCTGACATGAGAAACCGTACACTC 130

Qy        264 LeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnVal 283
              |||   ||||||||||||:::||||||||||||   ||||||||||||   |||||||||
Db        131 CTTCACAAACGACAATTTTTTCACTCACATAGAGCTCAGCCAATGGCAGCTGAGCAAGTT 190

Qy        284 MetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArg 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ATGTCAGATCGGGATAGTGAGGATGAAGTTGACGATGATGTTGCAGATTTTGAAGACCGA 250

Qy        304 GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrp 323
              :::|||||||||||||||||||||||   ||||||||||||||   ||||||||||||
Db        251 AGGATGCTTGATGATTTTGTAGACGTGACTAAAGATGAGAAGCAAATGATGCACTTGTGG 310

Qy        324 AsnSerPheValArg 328
              |||||||||||||||
Db        311 AACTCATTTGTGAGG 325

RESULT 14
AAC44147
ID   AAC44147 standard; DNA; 640 BP.
XX
AC   AAC44147;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 41806.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
```

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	1.29e-26	Length:	640
Score:	340.50	Matches:	67
Percent Similarity:	80.00%	Conservative:	1
Best Local Similarity:	78.82%	Mismatches:	1
Query Match:	14.32%	Indels:	16
DB:	21	Gaps:	2

US-09-890-220-2 (1-445) x AAC44147 (1-640)

```
Qy        362 CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIle 381
              |||||||||                   |||
Db        305 TGCTGGCGG-----------------TGG-------------------------TGC 319

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACCAAC 379

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 AACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGGAT 439

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        440 CATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGAA 499

Qy        442 -LysValIleLys 445
               ||||||||||||
Db        500 CAAAGTAATTAAA 512
```

RESULT 15
AAC51582
ID AAC51582 standard; DNA; 319 BP.
XX
AC AAC51582;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69042.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	1.25e-25	Length:	319
Score:	327.00	Matches:	61
Percent Similarity:	96.83%	Conservative:	0
Best Local Similarity:	96.83%	Mismatches:	2
Query Match:	13.75%	Indels:	0
DB:	21	Gaps:	0

US-09-890-220-2 (1-445) x AAC51582 (1-319)

```
Qy        383 AsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsn 402
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db          3 AACTGCAATACCATCCTCGAGAATTGGCGTAATAGCTCAGACACCACCACCACCAACAAC 62

Qy        403 AsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHis 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         63 AACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTGAGTCAT 122

Qy        423 ProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAspLys 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGAGACAAA 182

Qy        443 ValIleLys 445
              |||||||||
Db        183 GTAATTAAA 191
```

Search completed: June 19, 2003, 23:51:41
Job time : 377 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:39:38 ; Search time 70 Seconds
(without alignments)
1949.588 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144852_16502/app_query.fasta_1.583
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_40_@runat_13062003_144852_16502 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	120.5	5.1	198	5	PCT-US95-10668-3	Sequence 3, Appli
	2	120.5	5.1	198	5	PCT-US95-10668-4	Sequence 4, Appli
	3	114	4.8	198	5	PCT-US95-10668-1	Sequence 1, Appli
	4	114	4.8	198	5	PCT-US95-10668-2	Sequence 2, Appli
	5	109	4.6	2237	4	US-08-914-999-7	Sequence 7, Appli
	6	107	4.5	1086	1	US-08-415-751-27	Sequence 27, Appl
c	7	107	4.5	1086	1	US-08-415-751-28	Sequence 28, Appl
	8	107	4.5	5163	3	US-08-700-651-1	Sequence 1, Appli
	9	107	4.5	5163	3	US-08-928-361B-4	Sequence 4, Appli
	10	107	4.5	5318	3	US-08-700-651-2	Sequence 2, Appli
	11	107	4.5	5318	3	US-08-928-361B-3	Sequence 3, Appli
	12	107	4.5	5511	3	US-08-928-361B-2	Sequence 2, Appli
	13	107	4.5	7334	3	US-08-928-361B-1	Sequence 1, Appli
	14	105.5	4.4	1430	1	US-08-276-452A-25	Sequence 25, Appl
	15	105.5	4.4	1430	2	US-08-798-744-25	Sequence 25, Appl
	16	104.5	4.4	1680	2	US-08-700-152A-3	Sequence 3, Appli
	17	103	4.3	4223	4	US-09-541-782-5	Sequence 5, Appli
	18	103	4.3	4223	4	US-09-723-820-5	Sequence 5, Appli
	19	98.5	4.1	4757	4	US-09-572-191-1	Sequence 1, Appli
	20	98.5	4.1	4757	4	US-09-723-262-1	Sequence 1, Appli
	21	98.5	4.1	4757	4	US-09-723-219-1	Sequence 1, Appli
c	22	98.5	4.1	7425	4	US-09-453-702B-212	Sequence 212, App
c	23	98	4.1	1086	1	US-08-415-751-47	Sequence 47, Appl
	24	97	4.1	1690	1	US-08-276-452A-24	Sequence 24, Appl
	25	97	4.1	1690	2	US-08-798-744-24	Sequence 24, Appl
	26	96.5	4.1	2093	1	US-08-287-001A-1	Sequence 1, Appli
	27	96.5	4.1	2093	5	PCT-US95-09941-1	Sequence 1, Appli
	28	96	4.0	1956	4	US-08-559-896B-1	Sequence 1, Appli
	29	94	4.0	8920	2	US-08-446-855A-1	Sequence 1, Appli
	30	94	4.0	8920	4	US-09-150-741-1	Sequence 1, Appli
c	31	93	3.9	700	4	US-09-236-097-11	Sequence 11, Appl
	32	93	3.9	5720	4	US-09-442-100-1	Sequence 1, Appli
	33	91.5	3.8	2136	4	US-09-177-249-1	Sequence 1, Appli
	34	91.5	3.8	2315	4	US-09-061-769A-1	Sequence 1, Appli
	35	91.5	3.8	2322	4	US-09-061-769A-4	Sequence 4, Appli
	36	91.5	3.8	2327	4	US-09-061-769A-3	Sequence 3, Appli
	37	91.5	3.8	5361	4	US-08-973-462-2	Sequence 2, Appli
	38	91.5	3.8	6152	4	US-08-973-462-1	Sequence 1, Appli
	39	91	3.8	1631	3	US-09-118-319-1	Sequence 1, Appli
	40	91	3.8	2288	2	US-08-568-459A-5	Sequence 5, Appli
	41	91	3.8	2288	2	US-08-487-826B-5	Sequence 5, Appli
	42	91	3.8	2288	4	US-09-210-288-5	Sequence 5, Appli
	43	91	3.8	3800	4	US-08-965-762-13	Sequence 13, Appl
	44	91	3.8	3800	4	US-09-911-927-13	Sequence 13, Appl
c	45	91	3.8	3800	4	US-09-911-927-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
; Sequence 3, Application PC/TUS9510668
; GENERAL INFORMATION:
; APPLICANT: James Eberwine
; TITLE OF INVENTION: A Method of Sequencing Proteins by
; TITLE OF INVENTION: Epitope Ordering and Protein
; TITLE OF INVENTION: Restriction Mapping
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10668
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/294,133
; FILING DATE: August 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: PENN-0137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 3:

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-3

Alignment Scores:
Pred. No.:              4.91e-06        Length:         198
Score:                  120.50          Matches:        24
Percent Similarity:     65.08%          Conservative:   17
Best Local Similarity:  38.10%          Mismatches:     21
Query Match:            5.07%           Indels:         1
DB:                     5               Gaps:           1

US-09-890-220-2 (1-445) x PCT-US95-10668-3 (1-198)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||            |||   :::|||||||::::::            |||
Db          5 AACAACAACAACAACAACAACAACAACAACAAGAATTCCAACAACAACAACAACAACAAC 64

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||:::   :::       :::::::::|||   ||||||||||||      :::
Db         65 AACAACAACAAC---AACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 121

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::   |||:::   ||||||  ||||||   :::|||  :::|||||||:::   :::
Db        122 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 181

Qy        442 LysValIle 444
              |||::::::
Db        182 AAGCTGATG 190

RESULT 2
PCT-US95-10668-4
; Sequence 4, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-4

Alignment Scores:
Pred. No.:              4.91e-06        Length:         198
Score:                  120.50          Matches:        24
Percent Similarity:     65.08%          Conservative:   17
Best Local Similarity:  38.10%          Mismatches:     21
Query Match:            5.07%           Indels:         1
DB:                     5               Gaps:           1

US-09-890-220-2 (1-445) x PCT-US95-10668-4 (1-198)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||            |||   :::|||||||::::::            |||
Db          5 AACAACAACAACAACAACAACAACAACAACAACAAGAATTCCAACAACAACAACAACAACAAC 64

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||:::   :::       :::::::::|||   ||||||||||||      :::
Db         65 AACAACAACAAC---AACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 121

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::   |||:::   ||||||  ||||||   :::|||  :::|||||||:::   :::
Db        122 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 181

Qy        442 LysValIle 444
              |||::::::
Db        182 AAGCTGATG 190

RESULT 3
PCT-US95-10668-1
; Sequence 1, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
```

```
PCT-US95-10668-1

Alignment Scores:
Pred. No.:              3.05e-05        Length:         198
Score:                  114.00          Matches:        23
Percent Similarity:     60.00%          Conservative:   13
Best Local Similarity:  38.33%          Mismatches:     22
Query Match:            4.79%           Indels:         2
DB:                     5               Gaps:           1

US-09-890-220-2 (1-445) x PCT-US95-10668-1 (1-198)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||            |||      |||:::   |||             |||
Db          2 AACAACAACAACAAC------AACAACAACAACAACAACAAGGATCCCAACAACAACAAC 55

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||:::   :::      :::::::::|||   ||||||||||||      :::
Db         56 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 115

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::   |||:::   ||||||   ||||||   :::|||   :::|||||||:::   :::
Db        116 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 175

RESULT 4
PCT-US95-10668-2
; Sequence 2, Application PC/TUS9510668
;  GENERAL INFORMATION:
;    APPLICANT: James Eberwine
;    TITLE OF INVENTION:  A Method of Sequencing Proteins by
;    TITLE OF INVENTION:  Epitope Ordering and Protein
;    TITLE OF INVENTION:  Restriction Mapping
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10668
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/294,133
;      FILING DATE:  August 22, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
PCT-US95-10668-2

Alignment Scores:
Pred. No.:              3.05e-05        Length:         198
Score:                  114.00          Matches:        23
Percent Similarity:     60.00%          Conservative:   13
Best Local Similarity:  38.33%          Mismatches:     22
Query Match:            4.79%           Indels:         2
DB:                     5               Gaps:           1

US-09-890-220-2 (1-445) x PCT-US95-10668-2 (1-198)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||            |||      |||:::   |||             |||
Db          2 AACAACAACAACAAC------AACAACAACAACAACAACAAGGATCCCAACAACAACAAC 55

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||:::   :::      :::::::::|||   ||||||||||||      :::
Db         56 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 115

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::   |||:::   ||||||   ||||||   :::|||   :::|||||||:::   :::
Db        116 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 175

RESULT 5
US-08-914-999-7
; Sequence 7, Application US/08914999
; Patent No. 6346406
;  GENERAL INFORMATION:
;    APPLICANT:  Ryazanov, Alexey G.
;    APPLICANT:  Hait, William N.
;    APPLICANT:  Pavur, Karen S.
;    TITLE OF INVENTION:  ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
;    TITLE OF INVENTION:  AND METHODS OF USE THEREFOR
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David A. Jackson, Esq.
;      STREET:  411 Hackensack Ave, Continental Plaza, 4th
;      STREET:  Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/914,999
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  601-1-078
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2237 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Dictyostelium discoideum
US-08-914-999-7

Alignment Scores:
Pred. No.:              0.00736         Length:         2237
Score:                  109.00          Matches:        56
Percent Similarity:     43.69%          Conservative:   34
Best Local Similarity:  27.18%          Mismatches:     75
Query Match:            4.58%           Indels:         45
DB:                     4               Gaps:           10

US-09-890-220-2 (1-445) x US-08-914-999-7 (1-2237)
```

```
Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||   ||||||||||||      :::::::::|||   |||             |||
Db        620 CATTTAAACCTCCAAAAAAGATTGAATTTTTACAATCATGC-GTT-----------TTA 666

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluVal------AspAspAspValAla 298
              |||   |||      ||||||  |||  |||  :::   :::                
Db        667 GAGTTTGTA------GATAGAACATCATCAGATTTAATTTGTGGAGCAGAACCATATGTA 720

Qy        299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLys--- 317
              :::   :::   |||:::      ::::::         |||:::   ||||||:::  
Db        721 GAAGGACAATATAGAAAGTATAATAATAATAGTGGATTCGTTAGTAATGATGAAAGAAAT 780

Qy        318 ---------GlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAla 334
                      |||   |||                     |||   :::|||        
Db        781 ACACCACAATCATTCTCTCATTTCACATATGAACATTCAAATCATCAATTATTGATTATA 840

Qy        335 Asp------------GlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGlu 350
              |||             |||                     :::   :::        
Db        841 GATATTCAAGGTGTTGGT-------------------------GATCACTATACAGAC 873

Qy        351 LysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeu 370
              :::::::|||   |||           |||||||||   |||   |||:::   |||  
Db        874 CCACAAATTCATACCTATGA----------TGGTGTTGGTT-TTGGTATTGGTAAT-TTG 921

Qy        371 TrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu--- 389
                  |||                     :::   ||||||   |||   :::      
Db        922 GGTCAAAAAGGTTTTGAAAAGTTTTTAGATACTCATAAATGTAATGCAATTTGCCAATAT 981

Qy        390 ---AsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArg 408
                 |||   ::::::   :::      :::      ::::::      :::|||   |||  
Db        982 TTAAATTTACAATCAATTAATCCAAAATCTGAAAAAAGTGATTGTGGTACTGTACCAAGA 1041

Qy        409 ProSer--------------------------AspSerAsnThrAsnAsnAsnAsnIle 419
              |||                           |||:::|||   ||||||||||||      
Db       1042 CCAGATTTAATTTTCCCTGATACATCTGAAAGAGATAATAATAATAATAATAATAAT--- 1098

Qy        420 ValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
              ::::::   |||:::   ||||||   ||||||      |||   :::|||||:::  
Db       1099 ---AATAATAATAATAATAATAATAATAATAATAATAGTAATAATAATAATAATAAC 1155

Qy        440 ArgAspLysValIleLys 445
                    :::   |||      
Db       1156 AATAGTAGTATTTCAAAA 1173
```

```
RESULT 6
US-08-415-751-27
; Sequence 27, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1086 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
US-08-415-751-27

Alignment Scores:
Pred. No.:                0.00383        Length:          1086
Score:                    107.00         Matches:         23
Percent Similarity:       58.33%         Conservative:    12
Best Local Similarity:    38.33%         Mismatches:      25
Query Match:              4.50%          Indels:          0
DB:                       1              Gaps:            0

US-09-890-220-2 (1-445) x US-08-415-751-27 (1-1086)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||  |||           |||     |||::::::   |||      |||
Db        345 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 404

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              |||||||||         |||   :::   ::::::   ||||||||||||||||
Db        405 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 464

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::      :::   ||||||||||:::|||   :::|||   :::|||::::::   |||
Db        465 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 524

RESULT 7
US-08-415-751-28/c
; Sequence 28, Application US/08415751
; Patent No. 5643772
;  GENERAL INFORMATION:
;    APPLICANT:  PETERSEN, CAROLYN
;    APPLICANT:  LEECH, JAMES
;    APPLICANT:  NELSON, RICHARD, C.
;    APPLICANT:  GUT, JIRI
;    TITLE OF INVENTION:  POLYPEPTIDES BINDING ANTI-
;    TITLE OF INVENTION:  CRYPTOSPORIDIUM ANTIBODIES, DNA
;    TITLE OF INVENTION:  AND RNA ENCODING THEM, HYBRID
;    TITLE OF INVENTION:  VECTOR AND TRANSFORMED HOST AND
;    TITLE OF INVENTION:  METHODS FOR IMMUNOTHERAPY AND
;    TITLE OF INVENTION:  DIAGNOSIS AND KIT
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PHILLIPS, MOORE, LEMPIO & FINLEY
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Kb storage
;      COMPUTER:  PC
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/415,751
;      FILING DATE:  03-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/071,880
;      FILING DATE:  June 1, 1993
;      APPLICATION NUMBER:  07/891,301
;      FILING DATE:  May 29, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hana Dolezalova
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.19-2 (HHD)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-1677
;      TELEFAX:  (415) 324-1678
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1086 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Cryptosporidium parvum
US-08-415-751-28

Alignment Scores:
Pred. No.:              0.00383       Length:        1086
Score:                  107.00        Matches:       23
Percent Similarity:     58.33%        Conservative:  12
Best Local Similarity:  38.33%        Mismatches:    25
Query Match:            4.50%         Indels:        0
DB:                     1             Gaps:          0

US-09-890-220-2 (1-445) x US-08-415-751-28 (1-1086)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||           |||      |||:::::::      |||      |||
Db        742 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 683

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              |||||||||         |||   :::   ::::::   ||||||||||||||||
Db        682 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 623

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db        622 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 563

RESULT 8
US-08-700-651-1
; Sequence 1, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;  LENGTH: 5163
;  TYPE: DNA
;  ORGANISM: Cryptosporidium parvum
US-08-700-651-1

Alignment Scores:
Pred. No.:              0.0528        Length:        5163
Score:                  107.00        Matches:       23
Percent Similarity:     58.33%        Conservative:  12
Best Local Similarity:  38.33%        Mismatches:    25
Query Match:            4.50%         Indels:        0
DB:                     3             Gaps:          0

US-09-890-220-2 (1-445) x US-08-700-651-1 (1-5163)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||           |||      |||:::::::      |||      |||
Db       1812 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 1871

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              |||||||||         |||   :::   ::::::   ||||||||||||||||
Db       1872 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 1931

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db       1932 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 1991

RESULT 9
US-08-928-361B-4
; Sequence 4, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5163 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-4
```

```
Alignment Scores:
Pred. No.:                0.0528          Length:          5163
Score:                    107.00          Matches:           23
Percent Similarity:       58.33%          Conservative:      12
Best Local Similarity:    38.33%          Mismatches:        25
Query Match:              4.50%           Indels:             0
DB:                       3               Gaps:               0

US-09-890-220-2 (1-445) x US-08-928-361B-4 (1-5163)

Qy         382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
               ||||||  |||           |||      |||:::::::      |||      |||
Db        1812 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 1871

Qy         402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
               |||||||||         |||   :::   ::::::   ||||||||||||||||
Db        1872 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 1931

Qy         422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
               :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db        1932 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 1991

RESULT 10
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 5318
;   TYPE: DNA
;   ORGANISM: Cryptosporidium parvum
US-08-700-651-2

Alignment Scores:
Pred. No.:                0.0555          Length:          5318
Score:                    107.00          Matches:           23
Percent Similarity:       58.33%          Conservative:      12
Best Local Similarity:    38.33%          Mismatches:        25
Query Match:              4.50%           Indels:             0
DB:                       3               Gaps:               0

US-09-890-220-2 (1-445) x US-08-700-651-2 (1-5318)

Qy         382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
               ||||||  |||           |||      |||:::::       |||      |||
Db        1812 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 1871

Qy         402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
               |||||||||         |||   :::   ::::::   ||||||||||||||||
Db        1872 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 1931

Qy         422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
               :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db        1932 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 1991

RESULT 11
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-3

Alignment Scores:
Pred. No.:                0.0555          Length:          5318
Score:                    107.00          Matches:           23
Percent Similarity:       58.33%          Conservative:      12
Best Local Similarity:    38.33%          Mismatches:        25
Query Match:              4.50%           Indels:             0
DB:                       3               Gaps:               0

US-09-890-220-2 (1-445) x US-08-928-361B-3 (1-5318)

Qy         382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
               ||||||  |||           |||      |||:::::       |||      |||
Db        1812 AACAACCACAACAACTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 1871

Qy         402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
               |||||||||         |||   :::   ::::::   ||||||||||||||||
Db        1872 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 1931

Qy         422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
               :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db        1932 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 1991

RESULT 12
US-08-928-361B-2
; Sequence 2, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
```

```
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5511 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-2

Alignment Scores:
Pred. No.:              0.0589        Length:         5511
Score:                  107.00        Matches:        23
Percent Similarity:     58.33%        Conservative:   12
Best Local Similarity:  38.33%        Mismatches:     25
Query Match:            4.50%         Indels:         0
DB:                     3             Gaps:           0

US-09-890-220-2 (1-445) x US-08-928-361B-2 (1-5511)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||          |||      |||:::::::      |||      |||
Db       2163 AACAACCACAACAATTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 2222

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              |||||||||          |||   :::   ::::::   ||||||||||||||||
Db       2223 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 2282

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::      :::   |||||||||:::|||   :::|||  .:::|||::::::   |||
Db       2283 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 2342

RESULT 13
US-08-928-361B-1
; Sequence 1, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7334 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-1

Alignment Scores:
Pred. No.:              0.0953        Length:         7334
Score:                  107.00        Matches:        23
Percent Similarity:     58.33%        Conservative:   12
Best Local Similarity:  38.33%        Mismatches:     25
Query Match:            4.50%         Indels:         0
DB:                     3             Gaps:           0

US-09-890-220-2 (1-445) x US-08-928-361B-1 (1-7334)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||          |||      |||:::::::      |||      |||
Db       3831 AACAACCACAACAATTCTTAATCCAATCATTACAACAACAACTCAAAAACCAACAACAAC 3890

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              |||||||||          |||   :::   ::::::   ||||||||||||||||
Db       3891 AACAACAACAAAGGTTCCAGGTAAGCCACCAATAGCCACAACAACAACAACATTAAAGCC 3950

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              :::      :::   |||||||||:::|||   :::|||   :::|||::::::   |||
Db       3951 AATAGTTACAACAACAACAACAAAAGCAACAACAACAACAACAACAACAGTGCCAACGAC 4010

RESULT 14
US-08-276-452A-25
; Sequence 25, Application US/08276452A
; Patent No. 5646029
;  GENERAL INFORMATION:
;    APPLICANT:  Chen, Chao-Guang
;    APPLICANT:  Mau, Shaio-Lim
;    APPLICANT:  Du, He
;    APPLICANT:  Gane, Alison M
;    APPLICANT:  Bacic, Antony
;    APPLICANT:  Clarke, Adrienne E
;    TITLE OF INVENTION:  Plant Arabinogalactan Protein (AGP) Genes
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
```

```
;       ZIP:  80303
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/276,452A
;       FILING DATE:  18-JUL-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Caruthers, Jennie M.
;       REGISTRATION NUMBER:  34,464
;       REFERENCE/DOCKET NUMBER:  27-91A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (303)499-8080
;       TELEFAX:  (303)499-8089
;       TELEX:  49617824
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1430 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  2..1312
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  41..112
;       OTHER INFORMATION:  /note= "Derived amino acid sequence
;       OTHER INFORMATION:  corresponding to the peptide sequence by protein
;       OTHER INFORMATION:  microsequencing"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  41..112
;       OTHER INFORMATION:  /note= "Derived amino acids 14-24,
;       OTHER INFORMATION:  28, 30, 32-37 are identical to sequences obtained
;       OTHER INFORMATION:  by protein microsequencing"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  25..31
;       OTHER INFORMATION:  /note= "Amino acids 25, 27, 29, and
;       OTHER INFORMATION:  31 are hydroxylated proline residues; amino acid
;       OTHER INFORMATION:  26 can be T instead of A"
US-08-276-452A-25

Alignment Scores:
Pred. No.:                0.00927        Length:          1430
Score:                    105.50         Matches:         28
Percent Similarity:       42.11%         Conservative:    12
Best Local Similarity:    29.47%         Mismatches:      32
Query Match:              4.44%          Indels:          23
DB:                       1              Gaps:            2

US-09-890-220-2 (1-445) x US-08-276-452A-25 (1-1430)

Qy        372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
              |||:::|||      ::::::         ||||||   |||         ||||||
Db        551 AACAATGGCTACTCGGAGAATTACAACAACAACAACAACAATGGCTACTCCGAGAATTAC 610

Qy        392 ArgAsnSerSerAsp---------------------------------------------- 396
                  |||::::::::::
Db        611 AACAACAACAACAATGGCTACTCCAAGAATTACAACAACAATGGCTACTCCAAAAAAATC 670

Qy        397 ------------------ThrThrThrThrAsnAsnAsnAsnSerValAspArgProSer 410
                                 ||||||||||||
Db        671 AACAATAATGGTTACTCCCAGAATTACATGAACAACAACAACGGCTTCTCCGAGAGTTAC 730

Qy        411 AspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsn 430
              :::||||||   ||||||||||||||||          :::   ||||||   |||
Db        731 AACAGCAACAACAACAACAACAATATTTTCTCCGAGAATTACAACAACAATAATAACAAT 790

Qy        431 AsnVal------AspAsnLysAspAsnAsnSerArgAspLysVal 443
              ||||||      :::|||   :::|||||||:::   :::   |||
Db        791 AATGTTTTCTCCGAGAATTACAACAACAATAATAACAATAATGTT 835

RESULT 15
US-08-798-744-25
; Sequence 25, Application US/08798744
; Patent No. 5830747
;  GENERAL INFORMATION:
;    APPLICANT:  Chen, Chao-Guang
;    APPLICANT:  Mau, Shaio-Lim
;    APPLICANT:  Du, He
;    APPLICANT:  Gane, Alison M
;    APPLICANT:  Bacic, Antony
;    APPLICANT:  Clarke, Adrienne E
;    TITLE OF INVENTION:  Plant Arabinogalactan Protein (AGP) Genes
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  United States of America
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/798,744
;      FILING DATE:  13-FEB-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/276,452
;      FILING DATE:  18-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Caruthers, Jennie M.
;      REGISTRATION NUMBER:  34,464
;      REFERENCE/DOCKET NUMBER:  27-91A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303)499-8080
;      TELEFAX:  (303)499-8089
;      TELEX:  49617824
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1430 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2..1312
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  41..112
;      OTHER INFORMATION:  /note= "Derived amino acid sequence
;      OTHER INFORMATION:  corresponding to the peptide sequence by protein
;      OTHER INFORMATION:  microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  41..112
;      OTHER INFORMATION:  /note= "Derived amino acids 14-24,
;      OTHER INFORMATION:  28, 30, 32-37 are identical to sequences obtained
;      OTHER INFORMATION:  by protein microsequencing"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  25..31
;      OTHER INFORMATION:  /note= "Amino acids 25, 27, 29, and
;      OTHER INFORMATION:  31 are hydroxylated proline residues; amino acid
;      OTHER INFORMATION:  26 can be T instead of A"
```

US-08-798-744-25

Alignment Scores:

Pred. No.:	0.00927	Length:	1430
Score:	105.50	Matches:	28
Percent Similarity:	42.11%	Conservative:	12
Best Local Similarity:	29.47%	Mismatches:	32
Query Match:	4.44%	Indels:	23
DB:	2	Gaps:	2

US-09-890-220-2 (1-445) x US-08-798-744-25 (1-1430)

```
Qy        372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
              |||:::|||      ::::::         ||||||   |||         ||||||
Db        551 AACAATGGCTACTCGGAGAATTACAACAACAACAACAACAATGGCTACTCCGAGAATTAC 610

Qy        392 ArgAsnSerSerAsp---------------------------------------------- 396
                 |||:::::::::
Db        611 AACAACAACAACAATGGCTACTCCAAGAATTACAACAACAATGGCTACTCCAAAAAAATC 670

Qy        397 ------------------ThrThrThrThrAsnAsnAsnAsnSerValAspArgProSer 410
                                 ||||||||||||
Db        671 AACAATAATGGTTACTCCCAGAATTACATGAACAACAACAACGGCTTCTCCGAGAGTTAC 730

Qy        411 AspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsn 430
              :::||||||   ||||||||||||||         :::   ||||||   |||
Db        731 AACAGCAACAACAACAACAACAATATTTTCTCCGAGAATTACAACAACAATAATAACAAT 790

Qy        431 AsnVal------AspAsnLysAspAsnAsnSerArgAspLysVal 443
              ||||||      :::|||   :::|||||||:::   :::   |||
Db        791 AATGTTTTCTCCGAGAATTACAACAACAATAATAACAATAATGTT 835
```

Search completed: June 20, 2003, 01:25:08
Job time : 95 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 20, 2003, 00:55:28 ; Search time 231 Seconds
(without alignments)
2826.862 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144854_16586/app_query.fasta_1.583
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09890220_@CGN_1_1_156_@runat_13062003_144854_16586
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	423	17.8	2556	9	US-09-938-842A-2418	Sequence 2418, Ap
2	266	11.2	888	9	US-09-938-842A-453	Sequence 453, App
3	225	9.5	3545	10	US-09-764-864-144	Sequence 144, App
4	221	9.3	2217	9	US-09-874-162A-6	Sequence 6, Appli
5	221	9.3	2328	9	US-09-874-162A-9	Sequence 9, Appli
6	221	9.3	4409	9	US-09-874-162A-7	Sequence 7, Appli
7	221	9.3	4441	9	US-09-874-162A-4	Sequence 4, Appli
8	221	9.3	4441	10	US-09-969-347-169	Sequence 169, App
9	157.5	6.6	1257	10	US-09-822-830A-428	Sequence 428, App
10	153	6.4	2481	10	US-09-764-864-124	Sequence 124, App
c 11	129	5.4	4985	9	US-10-094-240-10	Sequence 10, Appl
12	128	5.4	2001	10	US-09-801-368-35	Sequence 35, Appl
13	127.5	5.4	2886	10	US-09-801-368-131	Sequence 131, App
14	124	5.2	3931	9	US-10-006-780-1	Sequence 1, Appli
c 15	123	5.2	264	10	US-09-969-373-472	Sequence 472, App
16	122.5	5.2	3996	9	US-10-087-464-42	Sequence 42, Appl
17	115	4.8	526	10	US-09-764-864-568	Sequence 568, App
c 18	113.5	4.8	32768	10	US-09-070-927A-100	Sequence 100, App
19	109	4.6	2237	9	US-09-832-292-11	Sequence 11, Appl
20	109	4.6	2237	10	US-09-994-485-7	Sequence 7, Appli
21	103	4.3	1299	9	US-09-938-842A-1872	Sequence 1872, Ap
22	103	4.3	3150	9	US-09-938-842A-243	Sequence 243, App
23	103	4.3	6868	9	US-10-228-531-13	Sequence 13, Appl
24	101.5	4.3	2713	9	US-09-981-876-129	Sequence 129, App
25	101.5	4.3	2713	9	US-09-148-545-129	Sequence 129, App
26	101	4.2	1605	10	US-09-801-368-123	Sequence 123, App
27	99.5	4.2	1145	10	US-09-764-864-582	Sequence 582, App
28	98.5	4.1	4775	9	US-10-146-473-81	Sequence 81, Appl
c 29	98.5	4.1	7425	9	US-10-114-170-212	Sequence 212, App
30	97	4.1	1320	9	US-10-083-357-593	Sequence 593, App
31	97	4.1	1680	10	US-09-801-368-255	Sequence 255, App
32	96.5	4.1	5085	9	US-10-198-846-9854	Sequence 9854, Ap
c 33	96.5	4.1	6723	10	US-09-070-927A-163	Sequence 163, App
34	96	4.0	333	9	US-09-938-842A-3152	Sequence 3152, Ap
35	96	4.0	1956	10	US-09-351-794A-1	Sequence 1, Appli
36	96	4.0	3282	10	US-09-801-368-391	Sequence 391, App
37	95	4.0	2098	9	US-10-286-264-91	Sequence 91, Appl
38	95	4.0	2277	10	US-09-801-368-223	Sequence 223, App
39	95	4.0	9279	10	US-09-801-368-171	Sequence 171, App
40	94.5	4.0	496	9	US-10-155-533-5	Sequence 5, Appli
41	94.5	4.0	951	9	US-10-155-533-10	Sequence 10, Appl
42	94.5	4.0	978	9	US-10-155-533-11	Sequence 11, Appl
43	94.5	4.0	978	9	US-10-155-533-12	Sequence 12, Appl
44	94.5	4.0	1672	9	US-10-153-668-75	Sequence 75, Appl
45	94.5	4.0	1678	9	US-10-155-533-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-938-842A-2418
; Sequence 2418, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2418
; LENGTH: 2556
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2418

Alignment Scores:

```
Pred. No.:              7.14e-41        Length:         2556
Score:                  423.00          Matches:        101
Percent Similarity:     40.39%          Conservative:   2
Best Local Similarity:  39.61%          Mismatches:     0
Query Match:            17.79%          Indels:         152
DB:                     9               Gaps:           2

US-09-890-220-2 (1-445) x US-09-938-842A-2418 (1-2556)

Qy         43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLys--- 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTAT 2160

Qy         61 ------------------------------------------------------------ 61

Db       2161 GCGTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGG 2220

Qy         62 ------------------------------------------------------SerArgSer 64
                                                                    |||||||||
Db       2221 AAGCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTCAAGATCT 2280

Qy         65 ThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGluVal 84
              ||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       2281 ACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAA--- 2337

Qy         85 ArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeuGln 104

Db       2337 ------------------------------------------------------------ 2337

Qy        105 PheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGluTyr 124

Db       2337 ------------------------------------------------------------ 2337

Qy        125 GlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGlySerAsp 144

Db       2337 ------------------------------------------------------------ 2337

Qy        145 AspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgArgGlnArgGly 164

Db       2337 ------------------------------------------------------------ 2337

Qy        165 GlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeu 184

Db       2337 ------------------------------------------------------------ 2337

Qy        185 ThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyr 204
                                                 ||||||||||||||||||||
Db       2338 -----------------------------------------GGAAACCGTGGTTTAGGATAT 2358

Qy        205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIle 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2359 CCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATTCCACCAGCCATA 2418

Qy        225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValPro 244
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db       2419 GCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAAGCTGTGGTCCCT 2478

Qy        245 AlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
              |||||||||||||||||||||||||||||||||||||||||||||
Db       2479 GCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGA 2523

RESULT 2
US-09-938-842A-453
; Sequence 453, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 453
;   LENGTH: 888
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-453

Alignment Scores:
Pred. No.:              1.93e-22        Length:         888
Score:                  266.00          Matches:        70
Percent Similarity:     55.73%          Conservative:   3
Best Local Similarity:  53.44%          Mismatches:     30
Query Match:            11.19%          Indels:         28
DB:                     9               Gaps:           4

US-09-890-220-2 (1-445) x US-09-938-842A-453 (1-888)

Qy        278 MetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspVal 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTT 60

Qy        298 AlaAspPheGluAspArgGln------------------------------------- 304
              |||||||||||||||||||
Db         61 GCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAGTAGGCAA 120

Qy        305 ------------------MetLeuAspAspPheValAspValAsnLysAspGluLysGln 318
                               ||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAAATGGTATACGATATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAA 180

Qy        319 PheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIle 338
              |||||||||||||||||||||||||||||||||||||||||          :::   ||||||
Db        181 TTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGA---------AACAAACACATC 231

Qy        339 SerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSer 358
                                  ||||||                  ||||||      |||
Db        232 TTATTACCAATCTTTGCTTTCATAGCGATCCCTCTCGCAGCCTTACAC------CTCTCC 285

Qy        359 LeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSer 378
              |||           |||       :::|||   ||||||
Db        286 TTA-----------ACTCTTACCTCCTTTCGCCTTAAAAACCATGTCTTTCGTTTAGAA 333

Qy        379 AlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
              |||     |||     :::|||     |||
Db        334 GCCTTAGCTAACGTTGTGCATACACGGTTTGAA 366

RESULT 3
US-09-764-864-144
; Sequence 144, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 144
;   LENGTH: 3545
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-864-144

Alignment Scores:
```

```
Pred. No.:              1.99e-16        Length:         3545
Score:                  225.00          Matches:        84
Percent Similarity:     37.09%          Conservative:   64
Best Local Similarity:  21.05%          Mismatches:     151
Query Match:            9.46%           Indels:         100
DB:                     10              Gaps:           12

US-09-890-220-2 (1-445) x US-09-764-864-144 (1-3545)

Qy         58 ArgLysArgLys--------------SerArgSerThrGlyMetValValPheAsnTyr 72
              ||||||   |||              :::|||           :::   :::   :::
Db        503 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 562

Qy         73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
                 |||||||||   |||::|||||   |||::|||      ||||||:::
Db        563 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 622

Qy         93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
              |||::::::   |||                |||      ||||||         |||
Db        623 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 682

Qy        113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
                 |||   |||   :::             :::         ::::::|||||:::
Db        683 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 739

Qy        133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152
                                  ||||||   |||                     |||
Db        740 --------------------AATGAGGGTTATGATGGC-----------------TCC 760

Qy        153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168
                    |||:::         ||||||   |||         ||||||
Db        761 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 820

Qy        169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
                 :::|||      :::     :::   :::      |||      |||
Db        821 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 880

Qy        188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
              :::   :::                  ||||||
Db        881 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG-------------------------- 913

Qy        208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db        913 ---------------------------------------------------------- 913

Qy        228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db        913 ---------------------------------------------------------- 913

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267
                 ::::::      :::|||:::   :::      |||               |||
Db        914 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC-----------AATCGT 955

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
              :::::::|||                |||:::      :::::::::      :::
Db        956 CTGTATTTCCATAGAGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1009

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
              ||||||||||||||   |||   :::             :::             ::::::
Db       1010 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1069

Qy        308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327
              :::|||   |||||||||:::   ||||||:::   |||   |||||||||      |||
Db       1070 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 1129

Qy        328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347
                 |||         |||||||||      ::::::   ||||||      |||
Db       1130 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 1189

Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
                 |||   :::::::::   :::      :::|||         :::   |||   |||
Db       1190 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 1246
```

```
Qy        368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
              :::   :::   :::      |||:::         :::|||:::         |||
Db       1247 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 1306

Qy        388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407
              |||         :::                                      ::::::
Db       1307 CTCCGTGAAATGCAG--------------------------------CAAAAATTAGAA 1333

Qy        408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424
              :::         ::::::      |||      |||   :::      |||
Db       1334 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 1393

Qy        425 ------------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
                          :::|||||||:::|||                  |||:::   |||
Db       1394 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 1450

RESULT 4
US-09-874-162A-6
; Sequence 6, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
;  APPLICANT: Koontz, Jason
;  APPLICANT:  Sklar, Jeffrey
;  TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
;  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
;  FILE REFERENCE: 05311-024001
;  CURRENT APPLICATION NUMBER: US/09/874,162A
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,093
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 2217
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-874-162A-6

Alignment Scores:
Pred. No.:              2.81e-16        Length:         2217
Score:                  221.00          Matches:        84
Percent Similarity:     37.09%          Conservative:   64
Best Local Similarity:  21.05%          Mismatches:     151
Query Match:            9.29%           Indels:         100
DB:                     9               Gaps:           12

US-09-890-220-2 (1-445) x US-09-874-162A-6 (1-2217)

Qy         58 ArgLysArgLys--------------SerArgSerThrGlyMetValValPheAsnTyr 72
              ||||||   |||              :::|||           :::   :::   :::
Db       1237 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 1296

Qy         73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
                 |||||||||   |||::|||||   |||::|||      ||||||:::
Db       1297 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 1356

Qy         93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
              |||::::::   |||                |||      ||||||         |||
Db       1357 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 1416

Qy        113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
                 |||   |||   :::             :::         ::::::|||||:::
Db       1417 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 1473

Qy        133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152
                                  |||      |||                     |||
Db       1474 --------------------AATGAGTGTTATGATGGC-----------------TCC 1494

Qy        153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168
                    |||:::         ||||||   |||         ||||||
Db       1495 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 1554

Qy        169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
```

```
Db       1555 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 1614

Qy        188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207

Db       1615 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG-------------------------- 1647

Qy        208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db       1647 ------------------------------------------------------------ 1647

Qy        228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db       1647 ------------------------------------------------------------ 1647

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267

Db       1648 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC-----------AATCGT 1689

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287

Db       1690 CTGTATTTCCATAGTGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1743

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307

Db       1744 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1803

Qy        308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327

Db       1804 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 1863

Qy        328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347

Db       1864 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 1923

Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367

Db       1924 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 1980

Qy        368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387

Db       1981 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 2040

Qy        388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407

Db       2041 CTCCGTGAAATGCAG--------------------------------CAAAAATTAGAA 2067

Qy        408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424

Db       2068 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 2127

Qy        425 ----------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439

Db       2128 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 2184
```

RESULT 5
US-09-874-162A-9
; Sequence 9, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: Koontz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
; TITLE OF INVENTION: ENDOMETRIAL STROMAL TUMORS
; FILE REFERENCE: 05311-024001
; CURRENT APPLICATION NUMBER: US/09/874,162A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,093
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-874-162A-9

Alignment Scores:

Pred. No.:	3.05e-16	Length:	2328
Score:	221.00	Matches:	84
Percent Similarity:	37.09%	Conservative:	64
Best Local Similarity:	21.05%	Mismatches:	151
Query Match:	9.29%	Indels:	100
DB:	9	Gaps:	12

US-09-890-220-2 (1-445) x US-09-874-162A-9 (1-2328)

```
Qy         58 ArgLysArgLys----------SerArgSerThrGlyMetValValPheAsnTyr 72

Db       1348 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 1407

Qy         73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92

Db       1408 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 1467

Qy         93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112

Db       1468 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 1527

Qy        113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132

Db       1528 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 1584

Qy        133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152

Db       1585 ----------------------AATGAGTGTTATGATGGC------------------TCC 1605

Qy        153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168

Db       1606 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 1665

Qy        169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187

Db       1666 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 1725

Qy        188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207

Db       1726 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG-------------------------- 1758

Qy        208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db       1758 ------------------------------------------------------------ 1758

Qy        228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db       1758 ------------------------------------------------------------ 1758

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267

Db       1759 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC-----------AATCGT 1800

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287

Db       1801 CTGTATTTCCATAGTGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1854

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307

Db       1855 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1914

Qy        308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327

Db       1915 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 1974

Qy        328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347

Db       1975 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 2034
```

```
Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
              |||   :::::::::   :::      :::|||         :::   |||   |||
Db       2035 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 2091

Qy        368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
              :::   :::   :::      |||:::         :::|||:::         |||
Db       2092 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 2151

Qy        388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407
              |||         :::                                     ::::::
Db       2152 CTCCGTGAAATGCAG-------------------------------CAAAAATTAGAA 2178

Qy        408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424
              :::         ::::::      |||      |||   :::      |||
Db       2179 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 2238

Qy        425 -----------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
                          :::|||||||:::|||                  |||:::   |||
Db       2239 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 2295

RESULT 6
US-09-874-162A-7
; Sequence 7, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
;  APPLICANT: Koontz, Jason
;  APPLICANT:  Sklar, Jeffrey
;  TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
;  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
;  FILE REFERENCE: 05311-024001
;  CURRENT APPLICATION NUMBER: US/09/874,162A
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,093
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 4409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52)...(2379)
US-09-874-162A-7

Alignment Scores:
Pred. No.:              8.83e-16      Length:         4409
Score:                  221.00        Matches:        84
Percent Similarity:     37.09%        Conservative:   64
Best Local Similarity:  21.05%        Mismatches:     151
Query Match:            9.29%         Indels:         100
DB:                     9             Gaps:           12

US-09-890-220-2 (1-445) x US-09-874-162A-7 (1-4409)

Qy         58 ArgLysArgLys--------------SerArgSerThrGlyMetValValPheAsnTyr 72
              ||||||   |||               :::|||           :::   :::   :::
Db       1399 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 1458

Qy         73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
                 |||||||||   |||:::||||||   |||:::|||       |||||||:::
Db       1459 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 1518

Qy         93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
              |||::::::   |||               |||      ||||||          |||
Db       1519 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 1578

Qy        113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
              |||   |||   :::          :::        ::::::||||||:::
Db       1579 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 1635

Qy        133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152
                                                  |||      |||                  |||
Db       1636 --------------------AATGAGTGTTATGATGGC------------------TCC 1656

Qy        153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168
                       |||:::         ||||||   |||           ||||||
Db       1657 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 1716

Qy        169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
                  :::|||      :::      :::  . :::      |||      |||
Db       1717 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 1776

Qy        188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
              :::   :::                  ||||||
Db       1777 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG------------------------- 1809

Qy        208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db       1809 ------------------------------------------------------------ 1809

Qy        228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db       1809 ------------------------------------------------------------ 1809

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267
                    ::::::      :::|||:::   :::      |||             |||
Db       1810 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC------------AATCGT 1851

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
                 ::::::||||||            |||:::     :::::::::      :::
Db       1852 CTGTATTTCCATAGTGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1905

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
              ||||||||||||||   |||   :::            :::             ::::::
Db       1906 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1965

Qy        308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327
              :::|||   ||||||||||:::   ||||||:::   |||   |||||||||       |||
Db       1966 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 2025

Qy        328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347
                 |||          |||||||||       ::::::   ||||||       |||
Db       2026 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 2085

Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
              |||   :::::::::   :::      :::|||         :::   |||   |||
Db       2086 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 2142

Qy        368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
              :::   :::   :::      |||:::         :::|||:::         |||
Db       2143 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 2202

Qy        388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407
              |||         :::                                     ::::::
Db       2203 CTCCGTGAAATGCAG-------------------------------CAAAAATTAGAA 2229

Qy        408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424
              :::         ::::::      |||      |||   :::      |||
Db       2230 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 2289

Qy        425 -----------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
                          :::|||||||:::|||                  |||:::   |||
Db       2290 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 2346

RESULT 7
US-09-874-162A-4
; Sequence 4, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
;  APPLICANT: Koontz, Jason
;  APPLICANT:  Sklar, Jeffrey
;  TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
;  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
;  FILE REFERENCE: 05311-024001
```

```
;  CURRENT APPLICATION NUMBER: US/09/874,162A
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,093
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 4441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (195)...(2411)
US-09-874-162A-4

Alignment Scores:
Pred. No.:                8.94e-16        Length:         4441
Score:                    221.00          Matches:        84
Percent Similarity:       37.09%          Conservative:   64
Best Local Similarity:    21.05%          Mismatches:     151
Query Match:              9.29%           Indels:         100
DB:                       9               Gaps:           12

US-09-890-220-2 (1-445) x US-09-874-162A-4 (1-4441)

Qy          58 ArgLysArgLys--------------SerArgSerThrGlyMetValValPheAsnTyr 72
               ||||||   |||                :::|||           :::   :::   :::
Db        1431 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 1490

Qy          73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
                  |||||||||   |||:::||||||   |||:::|||       ||||||:::
Db        1491 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 1550

Qy          93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
               |||::::::   |||                 |||       ||||||          |||
Db        1551 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 1610

Qy         113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
                  |||   |||   :::           :::          ::::::::::::::::
Db        1611 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 1667

Qy         133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152
                                    |||   |||                          |||
Db        1668 --------------------AATGAGTGTTATGATGGC-----------------TCC 1688

Qy         153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168
                        |||:::         ||||||   |||             ||||||
Db        1689 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 1748

Qy         169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
                  :::|||      :::       :::   :::       |||      |||
Db        1749 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 1808

Qy         188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
               :::   :::                  ||||||
Db        1809 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG-------------------------- 1841

Qy         208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db        1841 ------------------------------------------------------------ 1841

Qy         228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db        1841 ------------------------------------------------------------ 1841

Qy         248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267
                  ::::::      :::|||:::   :::       |||                  |||
Db        1842 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC------------AATCGT 1883

Qy         268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
                  ::::::||||||       |||:::      :::::::::      :::
Db        1884 CTGTATTTCCATAGTGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1937

Qy         288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
               ||||||||||||||||   |||   :::           :::           ::::::
Db        1938 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1997

Qy         308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327
                  :::|||   |||||||||:::   ||||||:::   |||   |||||||||      |||
Db        1998 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 2057

Qy         328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347
                  |||          |||||||||       ::::::   ||||||       |||
Db        2058 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 2117

Qy         348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
                  |||   :::::::::   :::    :::|||          :::   |||   |||
Db        2118 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 2174

Qy         368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
                  :::   :::   :::      |||:::          :::|||:::          |||
Db        2175 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 2234

Qy         388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407
                  |||          :::                                      ::::::
Db        2235 CTCCGTGAAATGCAG--------------------------------CAAAAATTAGAA 2261

Qy         408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424
                  :::          ::::::      |||      |||   :::      |||
Db        2262 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 2321

Qy         425 ------------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
                           :::|||||||:::|||                ||||::   |||
Db        2322 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 2378

RESULT 8
US-09-969-347-169
; Sequence 169, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-69
;  CURRENT APPLICATION NUMBER: US/09/969,347
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,598
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,604
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 318
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 169
;   LENGTH: 4441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-347-169

Alignment Scores:
Pred. No.:                8.94e-16        Length:         4441
Score:                    221.00          Matches:        84
Percent Similarity:       37.09%          Conservative:   64
Best Local Similarity:    21.05%          Mismatches:     151
Query Match:              9.29%           Indels:         100
DB:                       10              Gaps:           12

US-09-890-220-2 (1-445) x US-09-969-347-169 (1-4441)

Qy          58 ArgLysArgLys--------------SerArgSerThrGlyMetValValPheAsnTyr 72
               ||||||   |||                :::|||           :::   :::   :::
Db        1431 AGAAAAGAAAAGGATACTCCAAATGAAAACCGACAAAAATTAAGAATATTTTATCAGTTT 1490

Qy          73 LysAspCysAsnAsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
                  |||||||||   |||:::||||||   |||:::|||       ||||||:::
Db        1491 CTCTATAACAACAATACAAGGCAACAAACTGAAGCAAGAGATGACCTGCATTGCCCTTGG 1550
```

```
Qy         93 CysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
              |||::::::   |||               |||      ||||||         |||
Db       1551 TGTACTCTGAACTGCCGCAAACTTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 1610

Qy        113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
                 |||   |||   :::             :::         ::::::||||||:::
Db       1611 AGATTTATCTTCAACTATGTTTATCATCCAAAAGGTGCTAGGATAGATGTTTCTATC--- 1667

Qy        133 LeuAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSer 152
                          |||     |||                             |||
Db       1668 --------------------AATGAGTGTTATGATGGC-----------------TCC 1688

Qy        153 LeuCysSerLysProArgLys---ArgArgGlnArgGly---------GlyArgAsnAsn 168
                       |||:::        ||||||   |||           ||||||
Db       1689 TATGCAGGAAATCCTCAGGATATTCATCGCCAACCTGGATTTGCTTTTAGTCGCAACGGA 1748

Qy        169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
                 :::|||      :::      :::   :::      |||      |||
Db       1749 CCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCAAAACGAACAAAAGCA 1808

Qy        188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
              :::   :::             ||||||
Db       1809 AGCATGTCTGAATTTCTTGAATCTGAAGATGGG-------------------------- 1841

Qy        208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227

Db       1841 ---------------------------------------------------------- 1841

Qy        228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247

Db       1841 ---------------------------------------------------------- 1841

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267
                 ::::::     :::|||:::   :::      |||             |||
Db       1842 ------GAAGTAGAACAGCAAAGAACATATAGTAGTGGCCAC------------AATCGT 1883

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
                 ::::::||||||           |||:::      :::::::::      :::
Db       1884 CTGTATTTCCATAGTGATACCTGCTTACCTCTCCGTCCACAAGAAATG------GAAGTA 1937

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
              ||||||||||||||   |||   :::             :::             ::::::
Db       1938 GATAGTGAAGATGAAAAGGATCCTGAATGGCTAAGAGAAAAAACCATTACACAAATTGAA 1997

Qy        308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheVal 327
              :::|||  ||||||||||:::   ||||||:::   |||   |||||||||      |||
Db       1998 GAGTTTTCTGATGTTAATGAAGGAGAGAAAGAAGTGATGAAACTCTGGAATCTCCATGTC 2057

Qy        328 ArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArg 347
                 |||         |||||||||      ::::::   ||||||      |||
Db       2058 ATGAAGCATGGGTTTATTGCTGACAATCAAATGAATCATGCCTGTATGCTGTTTGTAGAA 2117

Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
                 |||   :::::::::   :::      :::|||         :::   |||   |||
Db       2118 AATTATGGACAGAAAATAATTAAG---AAGAATTTATGTCGAAACTTCATGCTTCATCTA 2174

Qy        368 IleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
              :::   :::   :::      |||:::      :::|||:::         |||
Db       2175 GTCAGCATGCATGACTTTAATCTTATTAGCATAATGTCAATAGATAAAGCTGTTACCAAG 2234

Qy        388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAsp 407
              |||         :::                                 ::::::
Db       2235 CTCCGTGAAATGCAG------------------------------CAAAAATTAGAA 2261

Qy        408 ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn--------- 424
              :::         ::::::      |||      |||   :::      |||
Db       2262 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 2321

Qy        425 ------------AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
                         :::|||||||:::|||                ||||:::   |||
Db       2322 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCA 2378

RESULT 9

US-09-822-830A-428
; Sequence 428, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
;  APPLICANT: Genetics Institute, Inc.
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Howes, Steven H.
;  APPLICANT:  Resnick, Richard J.
;  APPLICANT:  Gulukota, Kamalakar
;  APPLICANT:  Graham, James R.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
;  FILE REFERENCE: GIN 6402
;  CURRENT APPLICATION NUMBER: US/09/822,830A
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195,604
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 631
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 428
;   LENGTH: 1257
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 860
;   OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-428

Alignment Scores:
Pred. No.:              6.36e-09        Length:         1257
Score:                  157.50          Matches:        54
Percent Similarity:     45.65%          Conservative:   51
Best Local Similarity:  23.48%          Mismatches:     103
Query Match:            6.62%           Indels:         23
DB:                     10              Gaps:           5

US-09-890-220-2 (1-445) x US-09-822-830A-428 (1-1257)

Qy        217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIle 236
              |||      ::      ::::::|||      |||   ||||||   |||
Db        146 ACGGACCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCA-AAACGAACA 204

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
                  |||   :::            ::::::      ::::::      :::|||:::
Db        205 AAAGCAAGCATGTCTGAATTTCTTGAATCTGAAGATGGGGAAGTAGAACAGCAAAGAACA 264

Qy        257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
                 :::      |||              |||   ::::::||||||
Db        265 TATAGTAGTGGCCAC------------AATCGTCTGTATTTCCATAGTGATACCTGCTTA 312

Qy        277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
              |||:::      :::::::::      :::   ||||||||||||||   |||   :::
Db        313 CCTCTCCGTCCACAAGAAATG------GAAGTAGATAGTGAAGATGAAAAGGATCCTGAA 366

Qy        297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
                        :::            :::::::::|||   ||||||||:::   |||
Db        367 TGGCTAAGAGAAAAAACCATTACACAAATTGAAGAGTTTTCTGATGTTAATGAAGGAGAG 426

Qy        317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
              |||:::   |||   |||||||||      |||   |||         ||||||||
Db        427 AAAGAAGTGATGAAACTCTGGAATCTCCATGTCATGAAGCATGGGTTTATTGCTGACAAT 486

Qy        337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
                 ::::::   ||||||      |||         |||   :::::::::   :::
Db        487 CAAATGAATCATGCCTGTATGCTGTTTGTAGAAAATTATGGACAGAAAATAATTAAG--- 543

Qy        357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
                  :::|||         :::   |||   |||:::   :::   :::      |||:::
Db        544 AAGAATTTATGTCGAAACTTCATGCTTCATCTAGTCAGCATGCATGACTTTAATCTTATT 603
```

```
Qy        377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeu---------------------- 388
                  :::|||::        |||   |||
Db        604 AGCATAATGTCAATAGATAAAGCTGTTACCAAGCTCCGTGAAATGCAGCAAAAATTAGAA 663

Qy        389 ------GluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerVal 406
                  |||:::         :::::::::      |||         |||   :::
Db        664 AAGGGGGAATCTGCTTCCCCTGCAAACGAAGAAATAACTGAAGAACAAAATGGGACAGCA 723

Qy        407 AspArgProSerAspSerAsnThr---------------AsnAsnAsnAsnIleValAsp 421
                :::      |||:::   |||:::                    :::::::::
Db        724 AATGGATTTAGTGAAATTAACTCAAAAGAGAAAGCTTTGGAAACAGATAGTGTCTCAGGG 783

Qy        422 HisProAsnAspIleAsnAsnLysAsnAsn 431
                  |||         |||||||||:::
Db        784 TTTCAAAACAGAGCAAAAAACAAAAACTCT 813

RESULT 10
US-09-764-864-124
; Sequence 124, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 124
;   LENGTH: 2481
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-864-124

Alignment Scores:
Pred. No.:              7e-08          Length:          2481
Score:                  153.00         Matches:         53
Percent Similarity:     44.35%         Conservative:    49
Best Local Similarity:  23.04%         Mismatches:      103
Query Match:            6.43%          Indels:          26
DB:                     10             Gaps:            5

US-09-890-220-2 (1-445) x US-09-764-864-124 (1-2481)

Qy        217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIle 236
                |||     :::        ::::::|||      |||   ||||||   |||
Db        200 ACGGACCAGTTAAGAGAACACCTATCACACATATTCTTGTGTGCAGGCCA-AAACGAACA 258

Qy        237 LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
                  |||   :::          ::::::      ::::::        :::|||:::
Db        259 AAAGCAAGCATGTCTGAATTTCTTGAATCTGAAGATGGGGAAGTAGAACAGCAAAGAACA 318

Qy        257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
                 :::      |||               |||   ::::::||||||
Db        319 TATAGTAGTGGCCAC------------AATCGTCTGTATTTCCATAGTGATACCTGCTTA 366

Qy        277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
                |||:::      :::::::::      :::   ||||||||||||||   |||   :::
Db        367 CCTCTCCGTCCACAAGAAATG------GAAGTAGATAGTGAAGATGAAAAGGATCCTGAA 420

Qy        297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
                     :::              :::::::::|||   ||||||||||::   |||
Db        421 TGGCTAAGAGAAAAAACCATTACACAAATTGAAGAGTTTTCTGATGTTAATGAAGGAGAG 480

Qy        317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
                |||:::   |||   |||||||||      |||   |||          |||||||||
Db        481 AAAGAAGTGATGAAACTCTGGAATCTCCATGTCATGAAGCATGGGTTTATTGCTGACAAT 540

Qy        337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
                   ::::::   ||||||      |||         |||   :::::::::   ::::
Db        541 CAAATGAATCATGCCTGTATGCTGTTTGTAGAAAATTATGGACAGAAAATAATTAAG--- 597
```

```
Qy        357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
                 :::|||         :::   |||   |||:::   :::   :::      |||:::
Db        598 AAGAATTTATGTCGAAACTTCATGCTTCATCTAGTCAGCATGCATGACTTTAATCTTATT 657

Qy        377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAsp 396
                    :::|||:::        |||   |||        :::
Db        658 AGCATAATGTCAATAGATAAAGCTGTTACCAAGCTCCGTGAAATGCAG------------ 705

Qy        397 ThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsn 416
                                        :::::::::        ::::::      |||
Db        706 ---------------------CAAAAATTAGAAAAGGGGGAATCTGCTTCCCCTGCAAAC 744

Qy        417 AsnAsnIleValAspHisProAsn---------------------AspIleAsnAsnLys 429
                  |||   :::      |||                       :::||||||:::|||
Db        745 GAAGAAATAACTGAAGAACAAAATGGGACAGCAAATGGATTTAGTGAAATTAACTCAAAA 804

Qy        430 AsnAsnValAspAsnLysAspAsnAsnSer 439
                                  |||:::   |||
Db        805 GAGAAAGCTTTGGAAACAGATAGTGTCTCA 834

RESULT 11
US-10-094-240-10/c
; Sequence 10; Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
;  APPLICANT: ZWIEBEL, LAURENCE J.
;  TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
;  FILE REFERENCE: N8289
;  CURRENT APPLICATION NUMBER: US/10/094,240
;  CURRENT FILING DATE:  2001-03-08
;  PRIOR APPLICATION NUMBER: 10/056,405
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: 60/264,649
;  PRIOR FILING DATE: 2001-01-26
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-094-240-10

Alignment Scores:
Pred. No.:              0.000192       Length:          4985
Score:                  129.00         Matches:         33
Percent Similarity:     52.38%         Conservative:    22
Best Local Similarity:  31.43%         Mismatches:      32
Query Match:            5.42%          Indels:          18
DB:                     9              Gaps:            4

US-09-890-220-2 (1-445) x US-10-094-240-10 (1-4985)

Qy        348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
                |||   |||      ::::::   ||||||      :::|||   :::   ||||||
Db       3174 TTCTTAGAGATTATTGTGTATTACTACAGT------TACTGGAATTATTTACTATTTACT 3121

Qy        368 Ile-----------------LysLeuTrpAsnHisGlyLeuVal------------Asp 377
                |||                    ||||||      |||::::::
Db       3120 ATTTATTTACTTTATTATTTTTTATTATGGATTACTGGAATAATTAGAGAATAAAAAAAA 3061

Qy        378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
                :::      |||||||||                 |||      |||:::::::::
Db       3060 ACAATAATAATAAACAAT-----------------AATAAGAATAACAACAACAACAAT 3019

Qy        398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417
                          ||||||||||||:::            ::::::::::|||:::||||||
Db       3018 AATAAGAACAACAACAACAACAATAATAAGAATAATAATAATAACAATAGCAATAATAAG 2959

Qy        418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437
                |||      ::::::   |||:::   ||||||   ||||||   :::|||   :::|||
Db       2958 AATAATAATAACAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 2899

Qy        438 AsnSerArgAspLys 442
```

```
              |||::::::   |||
Db       2898 AATAATAAAAAAAAA 2884

RESULT 12
US-09-801-368-35
; Sequence 35, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
;   LENGTH: 2001
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-35

Alignment Scores:
Pred. No.:              5.58e-05      Length:          2001
Score:                  128.00        Matches:         86
Percent Similarity:     34.55%        Conservative:    56
Best Local Similarity:  20.92%        Mismatches:      118
Query Match:            5.38%         Indels:          152
DB:                     10            Gaps:            19

US-09-890-220-2 (1-445) x US-09-801-368-35 (1-2001)

Qy        118 PheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIle 137
              :::::::::      |||     |||:::  |||:::          |||:::
Db        109 TATGAAATAGAAAGACCTTATTTGATCGTATCAGGTAGTATA-----------TTCGTT 156

Qy        138 PheGluGluGluGlySer-----------------AspAspAspLysPheGluProPhe 151
              |||   ||||||  |||               |||         :::   |||
Db        157 TTTATTGAAGAGATATCGGGCATCAAGAGATGGACTGATGGCGTTTCTTGGTCTCCATCA 216

Qy        152 SerLeuCysSerLys-----------ProArgLysArgArgGlnArgGlyGlyArgAsn 167
                :::    |||                        ::::::      |||   |||
Db        217 AGAATATCAGGCAAGTTTCTAATATACAAGGAACTAGACAAGGAAAACGCCGGGTCTAAT 276

Qy        168 AsnThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
                                                                    |||
Db        277 ------------------------------------------------GCTAATGCT 285

Qy        188 ThrGluAsnGlyIleThr---------LeuLeuAsnAspGlyAsnArgGlyLeu---Gly 203
              |||   :::|||   |||         ::::::   ||||||      |||
Db        286 ACTTCGAGTGGCAGTACTGATTCCGCCGTCATTACGGATGGAACTTCTGGCGCGAGAAAT 345

Qy        204 TyrProGluAlaThrGluLeuAla-----------------GlyGlnPheGluMetThr 217
                |||   ::::::::::::                      ||||||:::
Db        346 AACCCTTCTTCGTCGAAAATCAAGCTACCTCCTTTGAAGAATCACCAGTTTGAT------ 399

Qy        218 SerAsnIleProProAlaIleAlaHisSerSerLeuAspAla----------------- 231
                  :::||||||   :::   |||||||||   ::::::
```

```
Db        400 ------TTGCCACCTACTATGGGCCATTCAAGTTTTGAGTCAGAGCAAGACACCTCCATT 453

Qy        232 ------------GlyAlaLysValIleLeuThrSerGluAlaValValProAlaThrLys 247
                         :::   :::   |||           :::|||      |||
Db        454 TCTCCATCAAACCGTTCGAACTTACCTTTAAAGTATACTGGCCTGGTGAAGAAGACAATA 513

Qy        248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArg 267
              :::   ||||||                                        ||||||
Db        514 TCTGTGAAGCTG-----------------------------------------AAGAGA 531

Qy        268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
                    ::::::|||     ::::::   :::   :::           |||   :::
Db        532 CCTCCATTCAACTCG------ATTGAAAATTTACATATTGTCTCTTATTATTCTGTAAAA 585

Qy        288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
              |||   :::   ::::::      |||         |||         :::
Db        586 GATAT-AAAACAAAATTGTCTAGTGACTCCTAAAGCTTCTCCTTTTTTGAAAGATGTTAG 644

Qy        308 AspPheValAspValAsn--------------------------LysAspGluLysGln 318
                |||         |||                       :::   |||||||||
Db        645 ACCTTCACAGGAACTAATTGTCGCGATGGGAAATACAACCTTAGGCAACGTGAAAAACAA 704

Qy        319 PheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIle 338
              |||   :::                                       |||
Db        705 TTCCACTAC-----------------------------------------CACCGG 719

Qy        339 SerTrpAla---------------------------------CysGluAlaPheSerArg 347
                 |||:::                                      :::
Db        720 TAATGGTCCTAATAATATAAACAATAAAAGCAACTCTTCCACTCCCCTGAATACTGTTAT 779

Qy        348 PheTyrGluLysGluLeuHisArgTyrSer---------------SerLeuPheTrpCys 362
              |||:::|||   :::         |||                 |||::::::   |||
Db        780 TTCCACGAATAACAATAGTGCTAATATAAATGCTGCTGGCAGTAATCAATTTACTAGTGC 839

Qy        363 TrpArgLeuPheLeuIle------------LysLeuTrp------AsnHisGlyLeuVal 376
                :::       |||:::           ||||||||      ||||||  :::
Db        840 AAACAAAAATTATTATTATAAAAATGATGAAAGCTCTGGATACCCAATCACTCAATTTGC 899

Qy        377 AspSerAlaThrIleAsnAsn---------------------------------------- 383
                      ||||||||||||
Db        900 TCCCGCATTACCATCAACAACTTTGATGTATACGGCGAATCCACCTTATATCACTCAATC 959

Qy        384 ----------------------------------Cys-AsnThrIleLeuGluAsnCy 391
                                                   ||| |||            |||
Db        960 TCCTGATAACACTAACGCTACTGGTATGAATACACATGTTAATAATAACAATAATAATAG 1019

Qy        391 sArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAs 411
                ||||||||:::::      :::||||||||||||
Db       1020 TAATAACAGTAGTAATAGTAATAATAGTAACAATAATAAT------------------- 1059

Qy        411 pSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAs 431
              |||   ||||||||||||      ::::::   |||:::|||||||||   |||||
Db       1060 ----AATAATAATAATAATAATAATAATAACAATAATAATAATATTAATAATATTAATAA 1115

Qy        431 nValAspAsnLysAspAsnAsnSerArgAsp 441
              ||||:::          |||     :::
Db       1116 TGTCAATACTAATGCAGGGAACGGTAACAAT 1146

RESULT 13
US-09-801-368-131
; Sequence 131, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
```

```
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 131
;   LENGTH: 2886
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-801-368-131

Alignment Scores:
Pred. No.:              0.000118      Length:          2886
Score:                  127.50        Matches:         26
Percent Similarity:     61.90%        Conservative:    13
Best Local Similarity:  41.27%        Mismatches:      23
Query Match:            5.36%         Indels:          1
DB:                     10            Gaps:            1

US-09-890-220-2 (1-445) x US-09-801-368-131 (1-2886)

Qy        382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsn 401
              ||||||   |||        :::|||      |||:::|||:::           |||
Db       1501 AACAACAACAATAACGATAACGATAACGATAACAATAATAGCAATAATAATAATAACAAT 1560

Qy        402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
              ||||||||||:::   :::      ::::::::::|||   ||||||||||||     :::
Db       1561 AATAATAATAACAATAATAATAATAACAATAATAATAACAATAACAACAACAACAACAAT 1620

Qy        422 HisProAsnAspIleAsnAsn---LysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              :::   |||:::   ||||||   ||||||||||||||||||   :::   |||
Db       1621 AATAATAATAATAGTAATAATATTAAAAATAACGTTGATAACAATAATACTAATCCAGCC 1680

Qy        441 AspLysVal 443
              |||   :::
Db       1681 GATAATATT 1689

RESULT 14
US-10-006-780-1
; Sequence 1, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
;  APPLICANT: Sakowicz, Roman
;  APPLICANT:  Beraud, Christophe
;  APPLICANT:  Guo, Jun
;  APPLICANT:  Freedman, Richard
;  TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
;  TITLE OF INVENTION:  METHODS FOR ITS USE
;  FILE REFERENCE: CYTOP083
;  CURRENT APPLICATION NUMBER: US/10/006,780
;  CURRENT FILING DATE:  2001-11-30
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3931
;   TYPE: DNA
;   ORGANISM: Plasmodium falciparum
US-10-006-780-1

Alignment Scores:
Pred. No.:              0.000529      Length:          3931
Score:                  124.00        Matches:         96
Percent Similarity:     36.36%        Conservative:    68
Best Local Similarity:  21.29%        Mismatches:      166
Query Match:            5.21%         Indels:          122
```

```
DB:                     9             Gaps:            21

US-09-890-220-2 (1-445) x US-10-006-780-1 (1-3931)

Qy         24 IleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeuGlyAsnPro 43
              |||||||||          :::||||||:::::          |||   :::
Db        483 ATTTATTGTGGT------AAATTATATGATTTATTACAAAAACGTAAGATG--------- 527

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSerArg 63
                                          :::   |||   ||||||   :::   |||
Db        528 ----------------------------GTAGCAGCATT-AGAAAATGGGAAAAAAGA 556

Qy         64 SerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu 83
              |||                 |||   |||      |||         ::::::
Db        557 AGTTGTAGTAAAAGATTTAAAAATATTAAGAGTATTAACAAA------AGAAGAATTAAT 610

Qy         84 ValArgGluAspCysSerCysProPheCysSerMetLeuCys-GlySerPheLysGlyLe 103
              ::::::  |||     |||                 |||  |||
Db        611 ATTAAAAATGATAGATGGTGTTTTATTAAGAAAAATTGGTGTTAATTCACAAAACGAT-- 668

Qy        103 uGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGl 123
                       :::   ||||||   :::          :::   |||
Db        669 -------GAATCATCTAGATCACATGCTATATTAAATATTGATTTAAAA----------- 710

Qy        123 uTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePhe-------------- 138
                     :::|||   :::      |||           |||
Db        711 -------GATATAAATAAAAATACATCTCTTGGAAAAATTGCTTTCATTGATTTAGCAGG 763

Qy        139 -GluGluGluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProAr 158
               |||   |||:::|||
Db        764 AAGTGAAAGAGGAGCTGAT-------------------------ACCGTTTCACAAAA 796

Qy        158 gLysArgArgGlnArgGlyGlyArgAsnAsnThrArg---------ArgLeuLysValCy 175
              |||:::   |||      |||  |||      |||           ||||||   ||
Db        797 TAAACAAACACAAACCGATGGAGCTAATATTAATAGATCTTTACTAGCCTTAAAGGAATG 856

Qy        175 sPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLe 195
              |         ::::||||||                    :::|||   |||
Db        857 TATTCGAGCTATGGATTCA-----------------GATAAAAATCATATACCTTTCAG 898

Qy        195 uAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu-------------- 210
              :::                          |||   |||::::::
Db        899 AGATTCA----------------------GAATTAACTAAAGTTTTAAGAGATATATT 934

Qy        211 -AlaGlyGlnPheGlu------MetThrSerAsnIleProProAlaIleAla-------- 225
               |||:::  :::      |||  :::||||||   |||   |||:::
Db        935 TGTAGGGAAATCTAAAAGTATTATGATAGCTAATATTTCTCCTACAATTAGTTGTTGTGA 994

Qy        226 ----------------------HisSerSerLeuAspAlaGlyAlaLysValIleLeuTh 238
                                   :::||||||                |||         ||
Db        995 GCAAACATTGAATACATTAAGATATTCTTCAAGAGTTAAGAACTTTAAAAATAAATCTAC 1054

Qy        238 rSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAl 258
              |                   |||   |||   :::::|||         |||:::
Db       1055 ATGTATAAATGAAGAAGATGATACAAATACCGAAAGAATTAGTATATTAGATTCAAAG-- 1112

Qy        258 aArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMe 278
                                                                  :::
Db       1113 ----------------------------------------GGAAGTGAAATGAATGCATC 1132

Qy        278 tAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAl 298
              ::::::|||   |||:::
Db       1133 TAGTATAGAAAATGTAGTT---------------------------------------- 1151

Qy        298 aAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGl 318
                    :::         :::|||   :::             |||   :::   |||
Db       1152 ----ATCAAATCAAATCATCTTCTTTCAAAT-----------AATAATAATAATAAAAT 1195

Qy        318 nPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIl 338
                          |||       :::   :::       ::::::   :::
Db       1196 CAATCGTGGCAAAATTAATGATAAGATAGAAAGAAATAATATCTTAAAAAATAAATCTTT 1255

Qy        338 eSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSe 358
```

```
                    |||   |||:::      :::   |||            ||||||||
Db       1256 TGATAAACCTAGGGAAGGATTCACATCAACCTTTGGAAAA-----------TACAGTTC 1303

Qy        358 rLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAsp-- 377
              ||||          ::::::         |||           ||||||::::::
Db       1304 ACTTAATGATATAGACAAAATAAAGAAAAATAAG------AAAAAGGGTTTAATTAATTA 1357

Qy        378 -SerAlaThrIleAsnAsnCysAsnThrIleLeuGlu-----------AsnCysArgAs 393
                :::|||:::   |||   |||||||||   :::            |||      ||
Db       1358 TAAAAGTACATTATATAATGACAACACCATAAATAAAAAGCACAACAATAATAATAATAA 1417

Qy        393 nSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAs 413
              |:::::::|||           ||||||:::|||:::   :::      :::|||||::
Db       1418 TAATAATGATAATAATGATAATAATAATGATAATAATAATAATAATAATAATGATAGTAG 1477

Qy        413 nThrAsnAsnAsnAsnIleValAspHis------ProAsnAspIleAsnAsnLysAsnAs 431
              ::::      ||||||:::::::::|||        |||:::|||||||||      ||
Db       1478 TAGTATGGTCAATAACATGATCAATCATATGATCAATAATAATATTAACAACAATATCAA 1537

Qy        431 nValAspAsnLysAspAsnAsnSerArgAsp 441
              ||||:::|||   :::||||||:::   :::
Db       1538 TGTAAATAACAACAATAATAATAATAATAAT 1568

RESULT 15
US-09-969-373-472/c
; Sequence 472, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 472
;   LENGTH: 264
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-472

Alignment Scores:
Pred. No.:              7.83e-06        Length:         264
Score:                  123.00          Matches:        26
Percent Similarity:     61.11%          Conservative:   18
Best Local Similarity:  36.11%          Mismatches:     24
Query Match:            5.17%           Indels:         4
DB:                     10              Gaps:           2

US-09-890-220-2 (1-445) x US-09-969-373-472 (1-264)

Qy        370 LeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
              :::|||:::         |||   :::   |||||||||           ||||||
Db        259 ATTTGGAGTTCT------GTCCCCAATTTGACTATAAATAAAAGTTTAAATATCCTC--- 209

Qy        390 AsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgPro 409
                  |||   |||:::::::::           ||||||||||||:::   :::
Db        208 ---TGTTTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 152

Qy        410 SerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLys 429
              :::::::::|||   ||||||||||||      ::::::   |||:::   ||||||
Db        151 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 92

Qy        430 AsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
              ||||||   :::|||   :::|||||||:::   :::
Db         91 AATAATAATAATAATAATAATAATAATAATAAT 56
```

Search completed: June 20, 2003, 02:35:54
Job time : 269 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:44:28 ; Search time 3000 Seconds
(without alignments)
3729.469 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144853_16527/app_query.fasta_1.583
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_2950_@runat_13062003_144853_16527 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2378	100.0	1722	33	US-09-890-220-1	Sequence 1, Appli
2	2378	100.0	1722	33	US-09-890-475-57	Sequence 57, Appl
3	2363	99.4	1722	33	US-09-890-220-47	Sequence 47, Appl
4	2285	96.1	1715	33	US-09-890-220-4	Sequence 4, Appli
5	2285	96.1	1715	33	US-09-890-475-58	Sequence 58, Appl
6	2258.5	95.0	1737	33	US-09-890-220-7	Sequence 7, Appli

	7	1923	80.9	1497	19	US-09-513-996A-69038	Sequence 69038, A
	8	1813.5	76.3	6338	33	US-09-890-220-3	Sequence 3, Appli
	9	1659	69.8	5895	33	US-09-890-220-6	Sequence 6, Appli
	10	1659	69.8	201471	20	US-09-534-859-587	Sequence 587, App
	11	1659	69.8	201471	31	US-09-803-736-587	Sequence 587, App
	12	1166	49.0	1722	42	US-10-219-999-5373	Sequence 5373, Ap
	13	923	38.8	2280	17	US-09-339-947A-2	Sequence 2, Appli
	14	907.5	38.2	2346	74	US-60-303-460-9	Sequence 9, Appli
	15	906.5	38.1	2248	17	US-09-339-947A-9	Sequence 9, Appli
	16	888.5	37.4	2249	74	US-60-303-460-1	Sequence 1, Appli
	17	871	36.6	2459	74	US-60-303-460-5	Sequence 5, Appli
	18	868.5	36.5	2276	74	US-60-303-460-7	Sequence 7, Appli
	19	833	35.0	1841	42	US-10-219-999-16339	Sequence 16339, A
	20	833	35.0	1841	75	US-60-312-544-4687	Sequence 4687, Ap
	21	833	35.0	1856	33	US-09-865-439A-20106	Sequence 20106, A
	22	833	35.0	2145	74	US-60-303-460-13	Sequence 13, Appl
	23	782.5	32.9	2296	74	US-60-303-460-11	Sequence 11, Appl
	24	654	27.5	2640	17	US-09-398-237-6	Sequence 6, Appli
	25	629	26.5	677	25	US-09-654-617-26623	Sequence 26623, A
	26	629	26.5	677	27	US-09-684-016-26623	Sequence 26623, A
	27	596.5	25.1	1176	74	US-60-303-460-3	Sequence 3, Appli
c	28	591.5	24.9	990	25	US-09-654-617-386601	Sequence 386601,
c	29	591.5	24.9	990	27	US-09-684-016-386601	Sequence 386601,
	30	559	23.5	704	74	US-60-303-460-17	Sequence 17, Appl
	31	527	22.2	600	43	US-10-260-238-2870	Sequence 2870, Ap
	32	527	22.2	600	61	US-60-170-912-179	Sequence 179, App
	33	527	22.2	600	76	US-60-325-448-2870	Sequence 2870, Ap
	34	521	21.9	492	43	US-10-260-238-1386	Sequence 1386, Ap
	35	521	21.9	492	76	US-60-325-448-1386	Sequence 1386, Ap
	36	510	21.4	726	43	US-10-260-238-2869	Sequence 2869, Ap
	37	510	21.4	726	76	US-60-325-448-2869	Sequence 2869, Ap
	38	492	20.7	2439	42	US-10-219-999-29250	Sequence 29250, A
	39	489	20.6	531	33	US-09-874-708A-29223	Sequence 29223, A
	40	489	20.6	531	65	US-60-211-750-29028	Sequence 29028, A
	41	472.5	19.9	443	63	US-60-197-872-14655	Sequence 14655, A
c	42	455	19.1	546	61	US-60-170-912-180	Sequence 180, App
	43	451.5	19.0	558	33	US-09-890-220-13	Sequence 13, Appl
c	44	451	19.0	453	61	US-60-172-375-387	Sequence 387, App
	45	443	18.6	566	33	US-09-865-439A-88732	Sequence 88732, A

ALIGNMENTS

```
RESULT 1
US-09-890-220-1
; Sequence 1, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-1

Alignment Scores:
Pred. No.:                3.72e-230      Length:          1722
Score:                    2378.00        Matches:         445
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       33             Gaps:            0

US-09-890-220-2 (1-445) x US-09-890-220-1 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              |||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 2
US-09-890-475-57
; Sequence 57, Application US/09890475
; GENERAL INFORMATION:
;  APPLICANT: Johanson, Urban
;  APPLICANT:  West, Joanne
;  APPLICANT:  Dean, Caroline
;  TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,475
;  CURRENT FILING DATE:  2001-11-13
;  PRIOR APPLICATION NUMBER: PCT/GB00/00197
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: GB 9902660.1
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 1722
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Landsberg
;   OTHER INFORMATION: erecta VRN2 cDNA
US-09-890-475-57

Alignment Scores:
Pred. No.:              3.72e-230       Length:         1722
Score:                  2378.00         Matches:        445
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     33              Gaps:           0

US-09-890-220-2 (1-445) x US-09-890-475-57 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550
```

```
Qy        441 AspLysValIleLys 445
              ||||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 3
US-09-890-220-47
; Sequence 47, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;    LENGTH: 1722
;    TYPE: DNA
;    ORGANISM: Arabidopsis thaliana
US-09-890-220-47

Alignment Scores:
Pred. No.:                1.24e-228      Length:          1722
Score:                    2363.00        Matches:         444
Percent Similarity:       99.78%         Conservative:    0
Best Local Similarity:    99.78%         Mismatches:      1
Query Match:              99.37%         Indels:          0
DB:                       33             Gaps:            0

US-09-890-220-2 (1-445) x US-09-890-220-47 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 290

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 350

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAACCCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGA 410

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAG 470

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 AAAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 530

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 TTCGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 650

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GAAGGAAGTGATGACGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGG 710

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 770

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGT 830

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACCAGCAACATT 890

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAA 950

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1010

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1070

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 GAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1130

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1190

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 CATCTTTGAAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1250

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTC 1310

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1370

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACCACCACCACC 1430

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 AACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1490

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 GATCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1550

Qy        441 AspLysValIleLys 445
              ||||||||||||||||
Db       1551 GACAAAGTAATTAAA 1565

RESULT 4
US-09-890-220-4
; Sequence 4, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics us
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 4
;   LENGTH: 1715
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-4

Alignment Scores:
Pred. No.:              1.01e-220      Length:         1715
Score:                  2285.00        Matches:        428
Percent Similarity:     97.51%         Conservative:   3
Best Local Similarity:  96.83%         Mismatches:     9
Query Match:            96.09%         Indels:         2
DB:                     33             Gaps:           1

US-09-890-220-2 (1-445) x US-09-890-220-4 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||   ||||||||||||||||||   ||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||   ||||||||||||||||||||||||||||   ||||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||:::|||         |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLys 442
              ||||||
Db       1547 GACAAG 1552

RESULT 5
US-09-890-475-58
; Sequence 58, Application US/09890475
; GENERAL INFORMATION:
;  APPLICANT: Johanson, Urban
;  APPLICANT:  West, Joanne
;  APPLICANT:  Dean, Caroline
;  TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late flow
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,475
;  CURRENT FILING DATE:  2001-11-13
;  PRIOR APPLICATION NUMBER: PCT/GB00/00197
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: GB 9902660.1
;  PRIOR FILING DATE: 1999-02-05
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
;   LENGTH: 1715
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
;   OTHER INFORMATION: cDNA
US-09-890-475-58

Alignment Scores:
Pred. No.:              1.01e-220      Length:         1715
Score:                  2285.00        Matches:        428
Percent Similarity:     97.51%         Conservative:   3
Best Local Similarity:  96.83%         Mismatches:     9
Query Match:            96.09%         Indels:         2
DB:                     33             Gaps:           1

US-09-890-220-2 (1-445) x US-09-890-475-58 (1-1715)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTCATATTTGAGGAA 652

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCTCTGCTCGAAACCTCGTAAGCGT 712

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 AGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGAT 772

Qy        181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
              ||||||||||||  ||||||||||||||||||||  ||||||||||||||||||||||||
Db        773 TCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCCCTGCTGAATGATGGAAACCGT 832

Qy        201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAACATT 892

Qy        221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        893 CCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAACCGAA 952

Qy        241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGC 1012

Qy        261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 CACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGCCAATGGCGCTT 1072

Qy        281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAGATTTT 1132

Qy        301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATG 1192

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTATAGCAGATGGTCATATCTCTTGG 1252

Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              |||||||||  ||||||||||||||||||||||||||||||||  ||||||||||||||||
Db       1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTC 1312

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGACTCAGCCACC 1372
```

```
Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              ||||||||||||||||||||||||||||||||||||||||:::|||          |||
Db       1373 ATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATACCTCA------GTCACTAAC 1426

Qy        401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1427 AACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACACCAACAACAATAACATTGTG 1486

Qy        421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1487 GATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAACAAGGACAATAACAGCAGA 1546

Qy        441 AspLys 442
              ||||||
Db       1547 GACAAG 1552
```

```
RESULT 6
US-09-890-220-7
; Sequence 7, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics us
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 1737
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-7

Alignment Scores:
Pred. No.:                5e-218          Length:          1737
Score:                    2258.50         Matches:         428
Percent Similarity:       95.78%          Conservative:    3
Best Local Similarity:    95.11%          Mismatches:      9
Query Match:              94.97%          Indels:          10
DB:                       33              Gaps:            2

US-09-890-220-2 (1-445) x US-09-890-220-7 (1-1737)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 292

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 352

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGA 412

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AAGTCAAGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAA 472

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGAACTGAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTC 532

Qy        101 Lys----------------------GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
              |||                      ||||||||||||||||||||||||||||||||||
Db        533 AAGGTGGGCAACTATTACAACTGAGGGGCTGCAATTTCATTTGAATTCATCTCATGATTT 592
```

```
Qy        113 uPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db        593 ATTTGAATTTGAGTTCAAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACT 652

Qy        133 uAsnSerPheIlePheGluGluGluGlySerAspAspAspLysPheGluProPheSerLe 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TAATTCCTTCATATTTGAGGAAGAAGGAAGTGATGATGATAAATTTGAGCCCTTCTCTCT 712

Qy        153 uCysSerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLy 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAA 772

Qy        173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh 193
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        773 AGTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGC 832

Qy        193 rLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGl 213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 CCTGCTGAATGATGGAAACCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACA 892

Qy        213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 ATTTGAGATGACTAGCAACATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGC 952

Qy        233 aLysValIleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAl 253
              ||||||||||||||::::||||||||||||||||||||||||||||||||||||||||||
Db        953 TAAAGTTATATTAACAACCGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGC 1012

Qy        253 aGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHi 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TGAGCGATCAGAGGCTAGAAGCCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCA 1072

Qy        273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVa 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 CAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGT 1132

Qy        293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 CGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGATGTGAA 1192

Qy        313 nLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIl 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 TAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGGGTTAT 1252

Qy        333 eAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLe 353
              ||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db       1253 AGCAGATGGTCATATCTCTTGGGCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTT 1312

Qy        353 uHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHi 373
              ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 GCACTGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCA 1372

Qy        373 sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAA 1432

Qy        393 nSerSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAs 413
              |:::|||         |||   ||||||||||||||||||   ||||||||||||||
Db       1433 TACCTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAA 1486

Qy        413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs 433
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1487 CACCAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGA 1546

Qy        433 pAsnLysAspAsnAsnSerArgAspLys 442
              ||||||||||||||||||||||||||||
Db       1547 CAACAAGGACAATAACAGCAGAGACAAG 1574

RESULT 7
US-09-513-996A-69038
; Sequence 69038, Application US/09513996A
; GENERAL INFORMATION:
;  APPLICANT: N. ALEXANDROV et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
;  TITLE OF INVENTION:  ENCODED THEREBY
;  FILE REFERENCE: 2750-709P
;  CURRENT APPLICATION NUMBER: US/09/513,996A
;  CURRENT FILING DATE:  2000-02-25
;  NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69038
;  LENGTH: 1497
;  TYPE: DNA
;  ORGANISM: Arabidopsis thaliana
;  FEATURE:
;  NAME/KEY: UNSURE
;  LOCATION: 1..1497
;  OTHER INFORMATION: any n or Xaa = unknown
;  FEATURE:
;  OTHER INFORMATION: Location 1..1497 / Ceres Seq. ID 2185972
US-09-513-996A-69038

Alignment Scores:
Pred. No.:                4.43e-184       Length:          1497
Score:                    1923.00         Matches:         370
Percent Similarity:       90.10%          Conservative:    3
Best Local Similarity:    89.37%          Mismatches:      7
Query Match:              80.87%          Indels:          34
DB:                       19              Gaps:            3

US-09-890-220-2 (1-445) x US-09-513-996A-69038 (1-1497)

Qy         43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 CCATCGTTTCTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTCA 381

Qy         63 ArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThr 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 AGATCTACTGGGATGGTAGTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACT 441

Qy         83 GluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGly 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GAAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCTATGTGGTAGCTTCAAGGGG 501

Qy        103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAG--------- 552

Qy        123 GluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGluGly 142
                 |||   :::||||||   ::::::         |||
Db        553 ---TACATCTCTGTAAATCCAACCATAATGTTTTGCAGC---------------------- 588

Qy        143 SerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgArgGln 162
                                                  ||||||||||||||||||||||
Db        589 ----------------------------------TCGAAACCTCGTAAGCGGAGACAA 612

Qy        163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTTTTTACCGTTGGATTCACCC 672

Qy        183 SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeu 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 AGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAATGATGGAAACCGTGGTTTA 732

Qy        203 GlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPro 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 GGATATCCCGAGGCAACAGAACTTGCTGGACAATTTGAGATGACCAGCAACATTCCACCA 792

Qy        223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVal 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGTG 852

Qy        243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg--------- 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 GTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGTACTGCC 912
```

```
Qy        260 ------------------------SerHisLeuLeuLeuGlnLysArgGlnPheTyrHis 271
                                      ||||||||||||||||||||||||||||||||||||
Db        913 ATATTAACCGAGAAAATTTCTTCCAGCCACCTACTTCTTCAGAAACGCCAATTCTATCAT 972

Qy        272 SerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAsp 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 TCTCACAGAGTCCAGCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGAT 1032

Qy        292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAsp 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 GAAGTCGATGACGATGTTGCAGATTTTGAAGATCGCCAGATGCTTGATGACTTTGTGGAT 1092

Qy        312 ValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArg 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 GTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTTGTAAGAAAACAAAGG 1152

Qy        332 ValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLys 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 GTTATAGCAGATGGTCATATTTCGTGGGCATGTGAAGCATTTTCAAGATTTTACGAGAAA 1212

Qy        352 GluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrp 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GAGTTGCACCGTTACTCATCACTCTTCTGGTGTTGGAGATTGTTTTTGATTAAACTATGG 1272

Qy        372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 AACCATGGACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGG 1332

Qy        392 ArgAsnSerSerAspThrThrThrAsnAsnAsnAsnSerValAspArgProSerAsp 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 CGTAATAGCTCAGACACCACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGAC 1392

Qy        412 SerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsn 431
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1393 TCAAACACCAACAACAATAACATTGTGAGTCATCCCAATGACATAAACAACAAGAACAAT 1452

Qy        432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
              |||||||||||||||||||||||||||||||||||||||||||
Db       1453 GTTGACAACAAGGACAATAACAGCAGAGACAAAGTAATTAAA 1494

RESULT 8
US-09-890-220-3
; Sequence 3, Application US/09890220
; GENERAL INFORMATION:
;  APPLICANT: Dean, Caroline
;  APPLICANT:  Gendall, Anthony
;  TITLE OF INVENTION: Methods and means for modification of plant characteristics using
;  TITLE OF INVENTION:  vernalisation gene VRN2.
;  FILE REFERENCE: Mewburn
;  CURRENT APPLICATION NUMBER: US/09/890,220
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: PCT/GB00/00248
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: GB 9901927.5
;  PRIOR FILING DATE: 1999-01-28
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 6338
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-3

Alignment Scores:
Pred. No.:              4.2e-172      Length:          6338
Score:                  1813.50       Matches:         440
Percent Similarity:     43.58%        Conservative:    1
Best Local Similarity:  43.48%        Mismatches:      4
Query Match:            76.26%        Indels:          569
DB:                     33            Gaps:            13

US-09-890-220-2 (1-445) x US-09-890-220-3 (1-6338)
```

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 2401

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2461

Qy         41 GlyAsn---------------------------------------------------- 42
              ||||||
Db       2462 GGCAACGTATGATTTGCCTTCCTCTCTCATCATATTAGCTCAGTAATCTTTCATCTCCTG 2521

Qy         42 ---------------------------------------------------------- 42

Db       2522 TGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGATTCATGGCGA 2581

Qy         43 ----------------------------------------------------ProSerPhe 45
                                                                   ||||||||
Db       2582 GTGTGTGCTTCTTTGTCTCCTAAATTTGAACTTGTTGTTTGTTGTTGCAGCCATCGTTT 2641

Qy         46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--------- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 CTTCCAAGATGCTTGAACTACAAAATTGGAGCAAAGCGCAAAAGAAAGTA-TGTTTTCTT 2700

Qy         62 ---------------------------------------------------------- 62

Db       2701 CTTGAATGTAGCTGCTACAGTGATATGTTATTTATCTTACTTCTAATATGGAAGCTGATG 2760

Qy         62 ---------------------------------------------------------- 62

Db       2761 ACCTATTTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTGTTCATG 2820

Qy         63 ----------------------------------------ArgSerThrGlyMetVal 68
                                                      ||||||||||||||||||
Db       2821 CTATAAACTTACATTTTATAAAATTGTGTTTTGGTTAGGTCAAGATCTACTGGGATGGTA 2880

Qy         69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------------- 83
              |||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GTTTTCAACTATAAGGATTGTAATAACACATTACAGAAAACTGAAGGTTAGTCTTTTTCT 2940

Qy         83 ---------------------------------------------------------- 83

Db       2941 GTTCGTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTATTTACTA 3000

Qy         84 ----------------------ValArgGluAspCysSerCysProPheCysSerMetLe 96
                                    ||||||||||||||||||||||||||||||||||||
Db       3001 TTTTTTTCTGTATTGTCATGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTCTATGCT 3060

Qy         96 uCysGlySerPheLys------------------------------------------- 101
              ||||||||||||||||
Db       3061 ATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCTTTCATA 3120

Qy        101 ---------------------------------------------------------- 101

Db       3121 TCTAACACTGTGAAATGCTACTGCTGTTTCATGCTGTATACTTTCACTGTTTGGTTACAT 3180

Qy        101 ---------------------------------------------------------- 101

Db       3181 ATTTTTGTGTTTGTTGTTTGTCTTCTCACTCTTTTCGAACTGCTGAGTGTGTGCTTATCT 3240

Qy        102 ----------------------------------------------------GlyLeuG 104
                                                                   |||||||
Db       3241 GAGAAAACATGTCCCAGATGGAGCTTACAACCAATTGTCTTGTGTCTATGCAGGGGCTGC 3300

Qy        104 lnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu---------- 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       3301 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATGTGGTTTTA 3360

Qy        120 ---------------------------------------------------------- 120

Db       3361 TGGAAATTCTTGATTTGCTATGCCTTTATTAATGAGGTTATAGTTAAAAAAGGGTCTTTC 3420

Qy        121 -----------Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
```

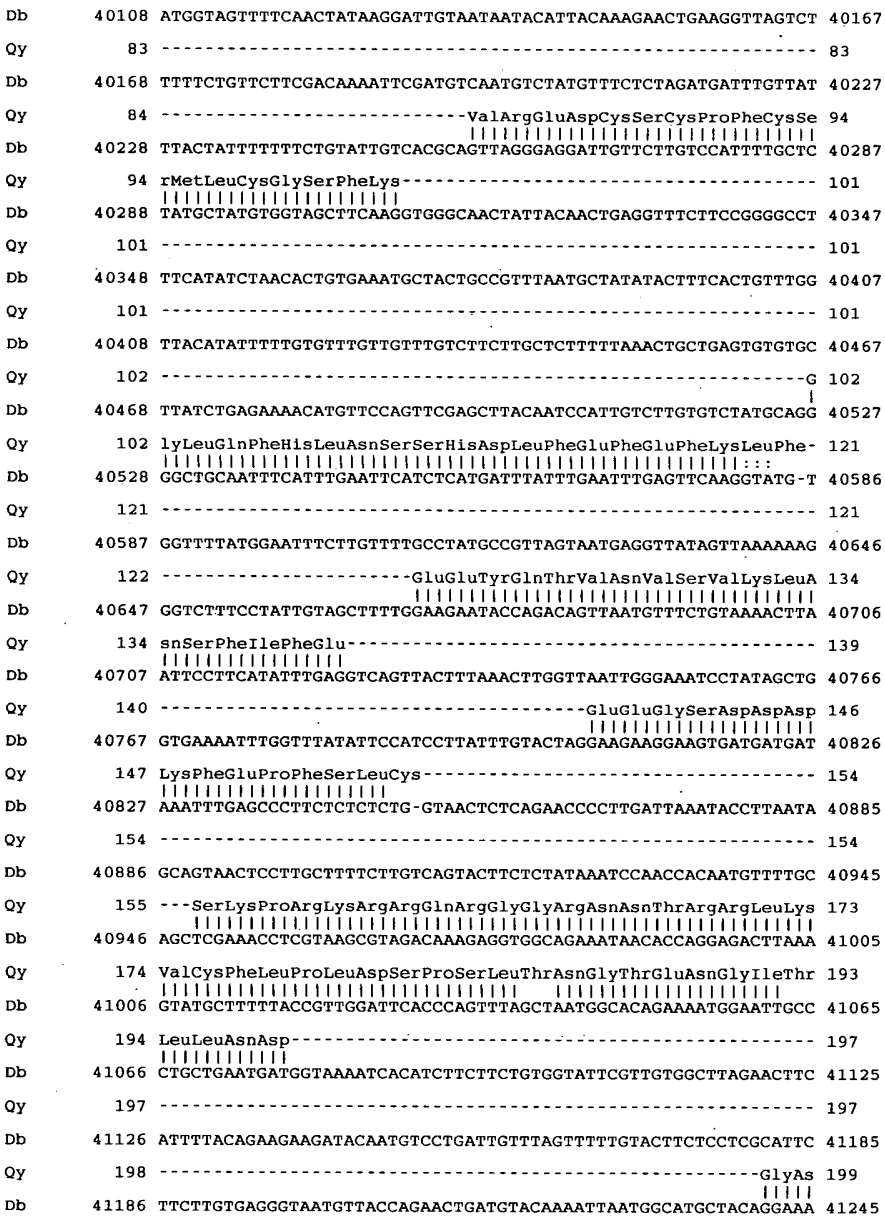
```
Db       3421 CTATTGTAGCTTTCGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTAATTCCTTC 3480

Qy        137 IlePheGlu---------------------------------------------------- 139
              |||||||||
Db       3481 ATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCGATAGCTGGTGAAAAT 3540

Qy        140 -----------------------------GluGluGlySerAspAspAspLysPheGl 149
                                            ||||||||||||||||||||||||||||||
Db       3541 TTTGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGACGATAAATTTGA 3600

Qy        149 uProPheSerLeu------------------------------------------------ 153
              ||||||||||||
Db       3601 GCCCTTCTCTCTCTGGTAACCCTCAGAACCCCTTCGATTAAATACCTTAATAGCAGTAAC 3660

Qy        154 ----------------------------------------------------CysSerLy 156
                                                                   ||||||
Db       3661 TCCTTGCTTCTCTTGTCAGTACATCTCTGTAAATCCAACCATAATGTTTTGCAGCTCGAA 3720

Qy        156 sProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPh 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ACCTCGTAAGCGGAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAAGTATGCTT 3780

Qy        176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAs 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TTTACCGTTGGATTCACCCAGTTTAACTAATGGCACAGAAAATGGAATCACCCTACTTAA 3840

Qy        196 nAsp---------------------------------------------------------- 197
              ||||
Db       3841 TGATGGTAAAATCATATCTTCTTCTGTGCGTTCCTTGTGGCTTAGAACTTCATATTACAG 3900

Qy        197 ---------------------------------------------------------- 197

Db       3901 AAGAAGATACAATGGCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTCTTCTTGCGA 3960

Qy        198 ----------------------------------------------GlyAsnArgGlyLe 202
                                                             ||||||||||||||
Db       3961 GGGTATTGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAACCGTGGTTT 4020

Qy        202 uGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPr 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AGGATATCCCGAGGCAACAGARCTTGCTGGACAATTTGAGATGACCAGCAACATTCCACC 4080

Qy        222 oAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVa 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 AGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTGACAAGCGAAGCTGT 4140

Qy        242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg-------- 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGAAGGTTTGT 4200

Qy        259 ---------------------------------------------------------- 259

Db       4201 TCATCATGACACCCCGTCATCATAATTACCATTCCTGTTGTTACAAATGTTCTTCCTATT 4260

Qy        260 ----------------------------------------------SerHisLeuLe 263
                                                           ||||||||||||
Db       4261 ATGGATAAGTGTTTATAGTACTGCCATATTAACCGAGAAAATTTCTTCCAGCCACCTACT 4320

Qy        263 uLeuGlnLysArgGlnPheTyrHisSerHisArgVal----------------------- 275
              ||||||||||||||||||||||||||||||||||
Db       4321 TCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGTTCCTTCACC 4380

Qy        275 ---------------------------------------------------------- 275

Db       4381 TACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGCATACTTGGT 4440

Qy        275 ---------------------------------------------------------- 275

Db       4441 TTGTTCTCATCCAAATTTGTATTTTGATCTGTATGTATCAACGCAAAATAGTTATGTCCA 4500

Qy        276 ----------------------------------------------Gln-ProMetA 279
                                                           ||| |||||||

Db       4501 TGTTGTCTCCGTTTTATTGCCACTAACCAAAAAATGCATGTTTCTGTGACAAGCCAATGG 4560

Qy        279 laLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaA 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 CGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGGATGAAGTCGATGACGATGTTGCAG 4620

Qy        299 spPheGluAspArg------------------------------------------------ 303
              |||||||||||||
Db       4621 ATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAATAGACAACAGA 4680

Qy        304 ----------------------------------------------------GlnMetLeu 306
                                                                   ||||||||
Db       4681 AAATGGTATATGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGCTGCAGATGCTT 4740

Qy        307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 GATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTTGGAACTCGTTT 4800

Qy        327 ValArgLysGlnArgVal------------------------------------------- 332
              |||||||||||||||
Db       4801 GTAAGAAAACAAAGGTA-ACTACTTCTCTTACACATGAACAGACACAAAAAGACCTTATG 4859

Qy        332 ---------------------------------------------------------- 332

Db       4860 TCTTACATTCCATACCTGTCTAAATGATTTTGCTTATGGAACTTTGAGCTCAATTATGAT 4919

Qy        333 ------------------IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346
                               ||||||||||||||||||||||||||||||||||||||||
Db       4920 TGTTGATGTTTCAGGGTTATAGCAGATGGTCATATTTCTTGGGCATGTGAAGCATTTTCA 4979

Qy        347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp--------------- 361
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       4980 AGATTTTACGAGAAAGAGTTGCACCGTTACTCATCACTCTTCTGGTAATATAAGTACACC 5039

Qy        361 ---------------------------------------------------------- 361

Db       5040 AAACATATACAGACACATAACTACACTATCAATCTTGTTTCGTTTTCTGAAAAAAAAATA 5099

Qy        362 ------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAsp 377
                           ||||||||||||||||||||||||||||||||||||||||||||||
Db       5100 AAAATTTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGGACTTGTCGAC 5159

Qy        378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5160 TCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAGCTCAGACACC 5219

Qy        398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5220 ACCACCACCAACAACAACAACAGTGTGGATCGTCCCAGTGACTCAAACACCAACAACAAT 5279

Qy        418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db       5280 AACATTGTGGRTCATCCCAATGACATAAACAACAAGAACAATGTTGACAACAAGGACAAT 5339

Qy        438 AsnSerArgAspLysValIleLys 445
              |||||||||||||||||||||||
Db       5340 AACAGCAGAGACAAAGTAATTAAA 5363
```

RESULT 9
US-09-890-220-6
; Sequence 6, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: Gendall, Anthony
; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
; TITLE OF INVENTION: vernalisation gene VRN2.
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; PRIOR FILING DATE: 1999-01-28

```
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 5895
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-890-220-6

Alignment Scores:
Pred. No.:              1.79e-156       Length:         5895
Score:                  1659.00         Matches:        424
Percent Similarity:     42.25%          Conservative:   4
Best Local Similarity:  41.86%          Mismatches:     12
Query Match:            69.76%          Indels:         577
DB:                     33              Gaps:           14

US-09-890-220-2 (1-445) x US-09-890-220-6 (1-5895)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1905 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 1964

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1965 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 2024

Qy         41 GlyAsn---------------------------------------------------------- 42
              |||
Db       2025 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 2083

Qy         42 ---------------------------------------------------------------- 42

Db       2084 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 2143

Qy         43 ----------------------------------------------------------Pro 43
                                                                        |||
Db       2144 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 2203

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2204 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 2262

Qy         62 ---------------------------------------------------------------- 62

Db       2263 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 2322

Qy         62 ---------------------------------------------------------------- 62

Db       2323 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 2382

Qy         63 ----------------------------------------------------ArgSerThrGly 66
                                                                  ||||||||||||
Db       2383 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 2442

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              ||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       2443 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 2502

Qy         83 ---------------------------------------------------------------- 83

Db       2503 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 2562

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           ||||||||||||||||||||||||||||||
Db       2563 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 2622

Qy         94 rMetLeuCysGlySerPheLys---------------------------------------- 101
              ||||||||||||||||||||
Db       2623 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 2682

Qy        101 ---------------------------------------------------------------- 101

Db       2683 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 2742

Qy        101 ---------------------------------------------------------------- 101

Db       2743 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 2802

Qy        102 ----------------------------------------------------------------G 102
                                                                               |
Db       2803 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 2862

Qy        102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       2863 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 2921

Qy        121 ---------------------------------------------------------------- 121

Db       2922 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 2981

Qy        122 ----------------------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
                                     ||||||||||||||||||||||||||||||||||||
Db       2982 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 3041

Qy        134 snSerPheIlePheGlu------------------------------------------- 139
              ||||||||||||||||
Db       3042 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 3101

Qy        140 ----------------------------------------GluGluGlySerAspAspAsp 146
                                                     ||||||||||||||||||||
Db       3102 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 3161

Qy        147 LysPheGluProPheSerLeuCys----------------------------------- 154
              ||||||||||||||||||||
Db       3162 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 3220

Qy        154 ---------------------------------------------------------------- 154

Db       3221 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 3280

Qy        155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3281 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 3340

Qy        174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db       3341 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 3400

Qy        194 LeuLeuAsnAsp------------------------------------------------- 197
              ||||||||||||
Db       3401 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 3460

Qy        197 ---------------------------------------------------------------- 197

Db       3461 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 3520

Qy        198 ----------------------------------------------------------GlyAs 199
                                                                         |||||
Db       3521 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 3580

Qy        199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3581 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 3640

Qy        219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db       3641 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 3700

Qy        239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  .     3701 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 3760

Qy        259 g----------------------------------------------------------- 259
              |
Db       3761 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 3820
```

```
Qy       260 ------------------------------------------------------------Se 260
                                                                          ||
Db      3821 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 3880

Qy       260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----------- 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3881 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 3940

Qy       276 ------------------------------------------------------------ 276

Db      3941 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 4000

Qy       276 ------------------------------------------------------------ 276

Db      4001 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 4060

Qy       276 ------------------------------------------------------------ 276

Db      4061 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 4120

Qy       277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4121 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 4180

Qy       295 pAspValAlaAspPheGluAspArg------------------------------------ 303
              |||||||||||||||||||||||
Db      4181 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 4240

Qy       303 ------------------------------------------------------------ 303

Db      4241 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 4300

Qy       304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4301 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 4360

Qy       323 rpAsnSerPheValArgLysGln-------------------------------------- 330
              |||||||||||||||||||||
Db      4361 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 4420

Qy       330 ------------------------------------------------------------ 330

Db      4421 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 4480

Qy       331 ---------------------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
                                     ||||||||||||||||||||||||||||||||||||
Db      4481 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 4540

Qy       344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
              ||||||||||||||||||||||||||||||   ||||||||||||||||||
Db      4541 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 4600

Qy       361,------------------------------------------------------------ 361

Db      4601 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 4660

Qy       362 ---------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
                                  ||||||||||||||||||||||||||||||||||||
Db      4661 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 4720

Qy       374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db      4721 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 4780

Qy       394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
              :|||         |||   ||||||||||||||||||   ||||||||||||||||||
Db      4781 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 4834

Qy       414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db      4835 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 4894

Qy       434 nLysAspAsnAsnSerArgAspLys 442
```

```
              ||||||||||||||||||||||||
Db      4895 CAAGGACAATAACAGCAGAGACAAG 4919

RESULT 10
US-09-534-859-587
; Sequence 587, Application US/09534859
; GENERAL INFORMATION:
;  APPLICANT:       Bush, David F.
;  APPLICANT:  Last, Robert L.
;  APPLICANT:  Levin, Irena M.
;  APPLICANT:  Norris, Susan R.
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Rounsley, Steven D.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION:       PLANT POLYMORPHIC MARKERS AND USES THEREOF
;  FILE REFERENCE:       38-10(15493)B
;  CURRENT APPLICATION NUMBER: US/09/534,859
;  CURRENT FILING DATE:  2000-03-29
;  NUMBER OF SEQ ID NOS:       1127
; SEQ ID NO 587
;   LENGTH: 201471
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-534-859-587

Alignment Scores:
Pred. No.:              2.39e-154       Length:         201471
Score:                  1659.00         Matches:        424
Percent Similarity:     42.25%          Conservative:   4
Best Local Similarity:  41.86%          Mismatches:     12
Query Match:            69.76%          Indels:         577
DB:                     20              Gaps:           14

US-09-890-220-2 (1-445) x US-09-534-859-587 (1-201471)

Qy         1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     39570 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 39629

Qy        21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     39630 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 39689

Qy        41 GlyAsn---------------------------------------------------------- 42
             |||
Db     39690 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 39748

Qy        42 ------------------------------------------------------------ 42

Db     39749 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 39808

Qy        43 ------------------------------------------------------------Pro 43
                                                                          |||
Db     39809 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 39868

Qy        44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     39869 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 39927

Qy        62 ------------------------------------------------------------ 62

Db     39928 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 39987

Qy        62 ------------------------------------------------------------ 62

Db     39988 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 40047

Qy        63 ----------------------------------------------ArgSerThrGly 66
                                                            ||||||||||||
Db     40048 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 40107

Qy        67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
             ||||||||||||||||||||||||||||||||||||||||:::|||||||
```


Db 40108 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 40167
Qy 83 ---------- 83
Db 40168 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 40227
Qy 84 ----------ValArgGluAspCysSerCysProPheCysSe 94
Db 40228 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 40287
Qy 94 rMetLeuCysGlySerPheLys---------- 101
Db 40288 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 40347
Qy 101 ---------- 101
Db 40348 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 40407
Qy 101 ---------- 101
Db 40408 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 40467
Qy 102 ----------G 102
Db 40468 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 40527
Qy 102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
Db 40528 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 40586
Qy 121 ---------- 121
Db 40587 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 40646
Qy 122 ----------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
Db 40647 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 40706
Qy 134 snSerPheIlePheGlu---------- 139
Db 40707 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 40766
Qy 140 ----------GluGluGlySerAspAspAsp 146
Db 40767 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 40826
Qy 147 LysPheGluProPheSerLeuCys---------- 154
Db 40827 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 40885
Qy 154 ---------- 154
Db 40886 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 40945
Qy 155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
Db 40946 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 41005
Qy 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
Db 41006 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 41065
Qy 194 LeuLeuAsnAsp---------- 197
Db 41066 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 41125
Qy 197 ---------- 197
Db 41126 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 41185
Qy 198 ----------GlyAs 199
Db 41186 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 41245
Qy 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
Db 41246 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 41305
Qy 219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
Db 41306 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 41365
Qy 239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
Db 41366 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 41425
Qy 259 g---------- 259
Db 41426 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 41485
Qy 260 ----------Se 260
Db 41486 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 41545
Qy 260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln---------- 276
Db 41546 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 41605
Qy 276 ---------- 276
Db 41606 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 41665
Qy 276 ---------- 276
Db 41666 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 41725
Qy 276 ---------- 276
Db 41726 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 41785
Qy 277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
Db 41786 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 41845
Qy 295 pAspValAlaAspPheGluAspArg---------- 303
Db 41846 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 41905
Qy 303 ---------- 303
Db 41906 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 41965
Qy 304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
Db 41966 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 42025
Qy 323 rpAsnSerPheValArgLysGln---------- 330
Db 42026 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 42085
Qy 330 ---------- 330
Db 42086 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 42145
Qy 331 ----------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
Db 42146 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 42205
Qy 344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
Db 42206 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 42265
Qy 361 ---------- 361
Db 42266 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 42325

```
Qy        362 ----------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
                                    ||||||||||||||||||||||||||||||||||||
Db      42326 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 42385

Qy        374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db      42386 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 42445

Qy        394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
              :|||        |||   ||||||||||||||||||||   ||||||||||||||||||
Db      42446 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 42499

Qy        414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db      42500 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 42559

Qy        434 nLysAspAsnAsnSerArgAspLys 442
              |||||||||||||||||||||||||
Db      42560 CAAGGACAATAACAGCAGAGACAAG 42584

RESULT 11
US-09-803-736-587
; Sequence 587, Application US/09803736
; GENERAL INFORMATION:
;  APPLICANT: Bush, David F.
;  APPLICANT:  Levin, Irena M.
;  APPLICANT:  Norris, Susan R.
;  APPLICANT:  Rounsley, Steven D.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
;  FILE REFERENCE: 38-10(15493)D
;  CURRENT APPLICATION NUMBER: US/09/803,736
;  CURRENT FILING DATE:  2001-03-12
;  PRIOR APPLICATION NUMBER: US 09/534,859
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: identified by Attorney Docket number 04983.0206CPUS01 38-10
;  PRIOR FILING DATE: 2000-10-20
;  NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 587
;   LENGTH: 201471
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-803-736-587

Alignment Scores:
Pred. No.:              2.39e-154     Length:          201471
Score:                  1659.00       Matches:         424
Percent Similarity:     42.25%        Conservative:    4
Best Local Similarity:  41.86%        Mismatches:      12
Query Match:            69.76%        Indels:          577
DB:                     31            Gaps:            14

US-09-890-220-2 (1-445) x US-09-803-736-587 (1-201471)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39570 ATGTGTAGGCAGAATTGTCGCGCGAAATCCTCACCGGAGGAAGTGATTTCAACTGATGAG 39629

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39630 AATCTCTTGATATATTGTAAACCTGTTCGACTATATAACATCTTTCACCTTCGCTCTCTA 39689

Qy         41 GlyAsn---------------------------------------------------- 42
              |||
Db      39690 GGCAA-CGTATGATTTGGCCTTCCTCTCTCATCATTTTAGCTTAGTAATCTTTCATCTCC 39748

Qy         42 ---------------------------------------------------------- 42

Db      39749 TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTATGGTCTGACTCATGGC 39808

Qy         43 ----------------------------------------------------------Pro 43
                                                                        |||
Db      39809 GAGTGTGTGCTTCTTTTGTCTCCTAATGTTATTTGAACTTGTTGTTTGTTGTTGCAGCCA 39868

Qy         44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39869 TCGTTTCTGCCAAGATGCTTGAACTACAAAATTGGGGCAAAGCGCAAAAGAAAGTA-TGC 39927

Qy         62 ---------------------------------------------------------- 62

Db      39928 GTTTCTTCTTGAATGTAGTTGCCACAGTGATATGTTATTTATCTTACTTCTAATATGGAA 39987

Qy         62 ---------------------------------------------------------- 62

Db      39988 GCTGATGAACTATTTATCTTTGTTGAGTAGATATGGACATAATGAATGGTTTCTTCTTTG 40047

Qy         63 ----------------------------------------------------ArgSerThrGly 66
                                                                   ||||||||||||
Db      40048 TTCATGCTATACACTTATATTTTACAAAATTGTGTTTTGCTTAGGTCAAGATCTACTGGG 40107

Qy         67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu--------- 83
              ||||||||||||||||||||||||||||||||||||||||:::||||||
Db      40108 ATGGTAGTTTTCAACTATAAGGATTGTAATAATACATTACAAAGAACTGAAGGTTAGTCT 40167

Qy         83 ---------------------------------------------------------- 83

Db      40168 TTTTCTGTTCTTCGACAAAATTCGATGTCAATGTCTATGTTTCTCTAGATGATTTGTTAT 40227

Qy         84 ----------------------------ValArgGluAspCysSerCysProPheCysSe 94
                                           ||||||||||||||||||||||||||||||||
Db      40228 TTACTATTTTTTTCTGTATTGTCACGCAGTTAGGGAGGATTGTTCTTGTCCATTTTGCTC 40287

Qy         94 rMetLeuCysGlySerPheLys---------------------------------- 101
              ||||||||||||||||||||||
Db      40288 TATGCTATGTGGTAGCTTCAAGGTGGGCAACTATTACAACTGAGGTTTCTTCCGGGGCCT 40347

Qy        101 ---------------------------------------------------------- 101

Db      40348 TTCATATCTAACACTGTGAAATGCTACTGCCGTTTAATGCTATATACTTTCACTGTTTGG 40407

Qy        101 ---------------------------------------------------------- 101

Db      40408 TTACATATTTTTGTGTTTGTTGTTTGTCTTCTTGCTCTTTTTAAACTGCTGAGTGTGTGC 40467

Qy        102 ----------------------------------------------------------G 102
                                                                           |
Db      40468 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACAATCCATTGTCTTGTGTCTATGCAGG 40527

Qy        102 lyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db      40528 GGCTGCAATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGTATG-T 40586

Qy        121 ---------------------------------------------------------- 121

Db      40587 GGTTTTATGGAATTTCTTGTTTTGCCTATGCCGTTAGTAATGAGGTTATAGTTAAAAAAG 40646

Qy        122 ----------------------GluGluTyrGlnThrValAsnValSerValLysLeuA 134
                                     ||||||||||||||||||||||||||||||||||||
Db      40647 GGTCTTTCCTATTGTAGCTTTTGGAAGAATACCAGACAGTTAATGTTTCTGTAAAACTTA 40706

Qy        134 snSerPheIlePheGlu---------------------------------------- 139
              ||||||||||||||||
Db      40707 ATTCCTTCATATTTGAGGTCAGTTACTTTAAACTTGGTTAATTGGGAAATCCTATAGCTG 40766

Qy        140 ------------------------------------GluGluGlySerAspAspAsp 146
                                                   ||||||||||||||||||||
Db      40767 GTGAAAATTTGGTTTATATTCCATCCTTATTTGTACTAGGAAGAAGGAAGTGATGATGAT 40826

Qy        147 LysPheGluProPheSerLeuCys---------------------------------- 154
              ||||||||||||||||||||
Db      40827 AAATTTGAGCCCTTCTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 40885

Qy        154 ---------------------------------------------------------- 154

Db      40886 GCAGTAACTCCTTGCTTTTCTTGTCAGTACTTCTCTATAAATCCAACCACAATGTTTTGC 40945
```

```
Qy       155 ---SerLysProArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys 173
Db     40946 AGCTCGAAACCTCGTAAGCGTAGACAAAGAGGTGGCAGAAATAACACCAGGAGACTTAAA 41005

Qy       174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
Db     41006 GTATGCTTTTTACCGTTGGATTCACCCAGTTTAGCTAATGGCACAGAAAATGGAATTGCC 41065

Qy       194 LeuLeuAsnAsp---------------------------------------------- 197
Db     41066 CTGCTGAATGATGGTAAAATCACATCTTCTTCTGTGGTATTCGTTGTGGCTTAGAACTTC 41125

Qy       197 ---------------------------------------------------------- 197
Db     41126 ATTTTACAGAAGAAGATACAATGTCCTGATTGTTTAGTTTTTGTACTTCTCCTCGCATTC 41185

Qy       198 ----------------------------------------------------GlyAs 199
Db     41186 TTCTTGTGAGGGTAATGTTACCAGAACTGATGTACAAAATTAATGGCATGCTACAGGAAA 41245

Qy       199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
Db     41246 CCGTGGTTTAGGATATCCCGAGGCAACAGAGCTTGCTGGACAATTTGAGATGACTAGCAA 41305

Qy       219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
Db     41306 CATTCCACCAGCCATAGCCCACTCTTCTCTGGACGCTGGTGCTAAAGTTATATTAACAAC 41365

Qy       239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
Db     41366 CGAAGCTGTGGTCCCTGCTACTAAGACAAGAAAGTTATCTGCTGAGCGATCAGAGGCTAG 41425

Qy       259 g---------------------------------------------------------- 259
Db     41426 AAGGTTTGTTCATCATGACACCCCGTCATCATAATTACCATACCTGTTGTTACAAATGTT 41485

Qy       260 ----------------------------------------------------------Se 260
Db     41486 CTTCCTATTATGGATAAGTGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCCAG 41545

Qy       260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----------- 276
Db     41546 CCACCTACTTCTTCAGAAACGCCAATTCTATCATTCTCACAGAGTCCAGGTGATCCAAGT 41605

Qy       276 ---------------------------------------------------------- 276
Db     41606 TCCTTCACCTACTTCTTAGGCATTTTCTTTAAATTGCTCATGATGATATCTTATCAAAGC 41665

Qy       276 ---------------------------------------------------------- 276
Db     41666 ATACTTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAAA 41725

Qy       276 ---------------------------------------------------------- 276
Db     41726 ATTATGTCCATGTTGTCTCCGTTTTATTGCCACTAACCAAAAACTGCATGTTTCTTGTGA 41785

Qy       277 ----ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
Db     41786 CAAGCCAATGGCGCTTGAGCAAGTAATGTCTGATCGGGATAGCGAGGATGAAGTCGATGA 41845

Qy       295 pAspValAlaAspPheGluAspArg----------------------------------- 303
Db     41846 CGATGTTGCAGATTTTGAAGATCGCCAGGTATTCCATGATTTCTTTCTGCGTTCATTAAG 41905

Qy       303 ---------------------------------------------------------- 303
Db     41906 TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTTGAAACTTAAAAAAGC 41965

Qy       304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
Db     41966 TGCAGATGCTTGATGACTTTGTGGATGTGAATAAAGATGAAAAGCAATTCATGCATCTTT 42025

Qy       323 rpAsnSerPheValArgLysGln------------------------------------ 330
```

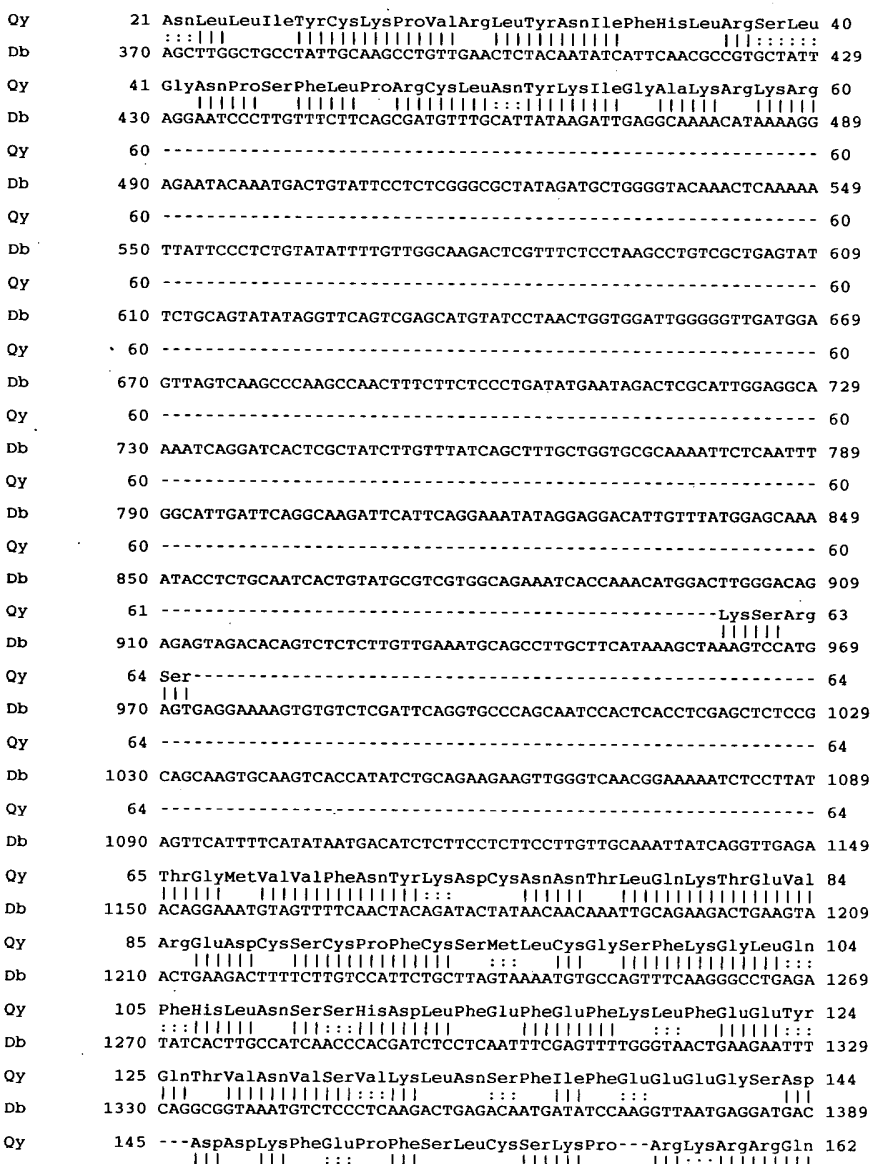
```
Db     42026 GGAACTCGTTTGTAAGAAAACAAAGGTAACTACTTCTCTTACACTTGAACACACACAAAA 42085

Qy       330 ---------------------------------------------------------- 330
Db     42086 AGACCTTATGTCTTACATTCCATACCTGTCTAAATGATTCTGCTTATGGAACTTTGAGCT 42145

Qy       331 ----------------------ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
Db     42146 CAAATTATGATTGATGTTTGCAGGGTTATAGCAGATGGTCATATCTCTTGGGCATGTGAA 42205

Qy       344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp------ 361
Db     42206 GTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTTACTCATCACTCTTCTGGTAATAT 42265

Qy       361 ---------------------------------------------------------- 361
Db     42266 AAGTACACCAAACATATACAGACACATAACTACACTATCAATTTTGTTTCGTTTTTCTGA 42325

Qy       362 ----------------------CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGl 374
Db     42326 AAGAAAAATAAAAAATTCCAGGTGTTGGAGATTGTTTTTGATTAAACTATGGAACCATGG 42385

Qy       374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
Db     42386 ACTTGTCGACTCAGCCACCATCAACAACTGCAATACCATCCTCGAGAATTGCCGTAATAC 42445

Qy       394 rSerAspThrThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
Db     42446 CTCA------GTCACTAACAACAACAACAACAGTGTGGATCATCCCAGTGACTCAAACAC 42499

Qy       414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
Db     42500 CAACAACAATAACATTGTGGATCATCCGAATGACATAAAAAACAAGAACAATGTTGACAA 42559

Qy       434 nLysAspAsnAsnSerArgAspLys 442
Db     42560 CAAGGACAATAACAGCAGAGACAAG 42584
```

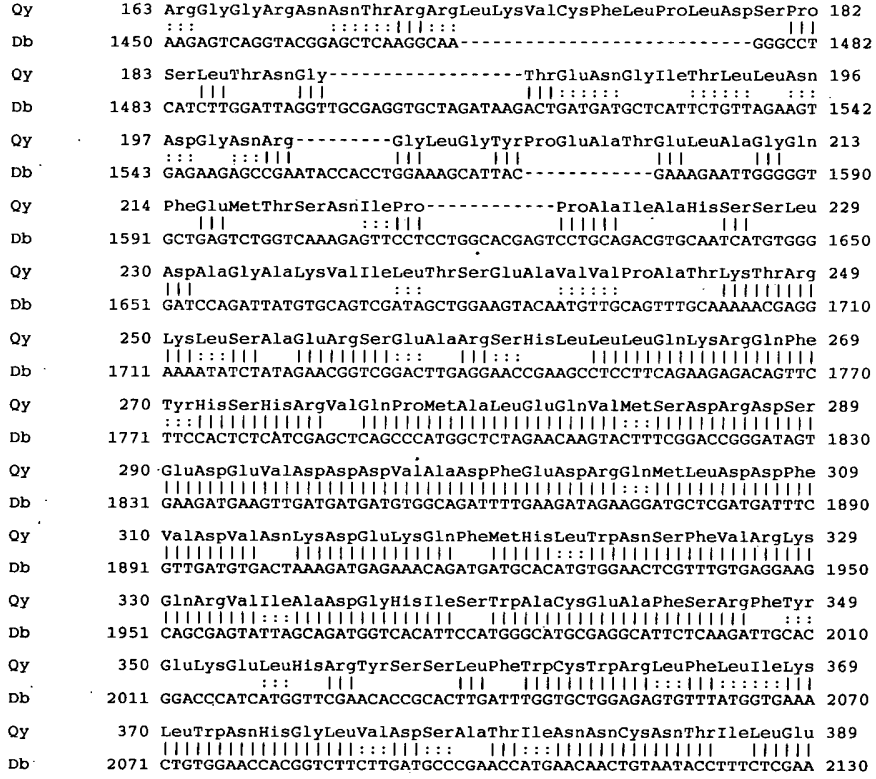
RESULT 12
US-10-219-999-5373
; Sequence 5373, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 5373
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1450)
; OTHER INFORMATION:
US-10-219-999-5373

Alignment Scores:

Pred. No.:	3.37e-107	Length:	1722
Score:	1166.00	Matches:	242
Percent Similarity:	65.65%	Conservative:	58
Best Local Similarity:	52.95%	Mismatches:	111

```
Query Match:          49.03%      Indels:        46
DB:                   42          Gaps:          9

US-09-890-220-2 (1-445) x US-10-219-999-5373 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||||||||              :::   ||||||   |||:::   ||||||
Db        164 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 223

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        224 AGCCTTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCTCTT 283

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||   |||||||||
Db        284 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAAAGG 343

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              :::   |||:::   |||:::|||:::|||||||||:::|||     ||||   |||:::
Db        344 AGGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTTCGA 400

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||||   |||   |||||||:::|||
Db        401 AAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAACTTT 460

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||:::|||||||||   |||||||||||||||||   |||||||||   :::
Db        461 AAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGGGTT 520

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
                 |||:::||||||   |||||||||||||||||||::::::      |||
Db        521 ACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGAT------ATATTGAGATCA 574

Qy        141 GluGlySerAspAsp---------AspLysPheGluProPheSerLeuCysSerLysPro 157
              |||         |||                :::   :::   |||      ||||||:::|||
Db        575 GAGAATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCT 634

Qy        158 ArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
              ||||||||||||::::::                  :::|||      |||   ||||||
Db        635 CGAAAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTG 694

Qy        178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197
                 ||||||||||||         |||      ||||||                :::|||
Db        695 GAGTTGGACTCACCA---------GAAGGCATACATAATGGATTTCTACAAAAAGATGAT 745

Qy        197 ------------------------------------------------------------ 197

Db        746 GATATCTTATCCTGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTC 805

Qy        198 --------------------GlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu 210
                                 |||      |||      :::|||          :::
Db        806 CCTAGTGGAAGAAATGACGGAGGAAAATTTGGTCCTGATCATCCTGGCACCATGGACAAC 865

Qy        211 AlaGlyGlnPheGluMetThrSerAsnIleProPro---------AlaIleAlaHisSer 227
                    |||   :::   |||||||||           |||:::      |||
Db        866 CTGGAGCATGTGGAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCT 925

Qy        228 SerLeuAspAlaGly---AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246
              |||:::|||         :::|||   |||      |||:::      :::||||||
Db        926 TCTGTGGACCCTGAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCT--- 982

Qy        247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLys 266
              ||||||:::|||||||||   :::||||||:::::::|||:::   :::||||||||||||
Db        983 AAAACAAAGAAGCTAAGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAG 1042

Qy        267 ArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAsp 286
              |||   |||:::|||||||||||||||||||||||||||||||||||||:::||||||
Db       1043 AGACTTTTCTTTCACTCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGAC 1102

Qy        287 ArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeu 306
              ||||||||||||||||||||||||||||||:::|||||||   ||||||||:::||||||
Db       1103 CGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTT 1162

Qy        307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
              ||||||||||||||||||:::|||||||||||||||   ||||||||||||||||||||
Db       1163 GACGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTT 1222

Qy        327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346
              :::||||||||||||||||:::|||||||||||||:::   ||||||||||||||||||||
Db       1223 ATGAGGAAGCAAAGGGTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCC 1282

Qy        347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPhe 366
              :::   :::   |||||||||           :::||||||||||||||||||||||||
Db       1283 AAGCTTCATGGAAAAGAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTC 1342

Qy        367 LeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
              :::|||||||||||||||||||||||||||:::|||:::   |||:::||||||||:::
Db       1343 ATGATCAAACTTTGGAATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATA 1402

Qy        387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerVal 406
              :::|||::::::   ||||||             |||      |||
Db       1403 GTATTAGATAGTTACAGAAATGAGGGATCGGGTACAAGAAAAAAT--------------- 1447

Qy        407 AspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisPro 423
                 ||||||      :::      ||||||   ||||||         |||
Db       1448 TGAAGACCATAAGGAAACTGACTCAACAACAAGAATATACAGTCTTCTCCG 1498

RESULT 13
US-09-339-947A-2
; Sequence 2, Application US/09339947A
; GENERAL INFORMATION:
;  APPLICANT:  YOSHIDA, Nobumasa
;  APPLICANT:  YANAI, Yukihiro
;  APPLICANT:  KATO, Yoshihiro
;  APPLICANT:  HIRATSUKA, Junzo
;  APPLICANT:  TAKAHASHI, Shigeru
;  APPLICANT:  MIWA, Tatsushi
;  TITLE OF INVENTION:  GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
;  TITLE OF INVENTION:  FLOWERING
;  FILE REFERENCE:  032735-001
;  CURRENT APPLICATION NUMBER: US/09/339,947A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:  JP 1998-180065
;  PRIOR FILING DATE:  1998-06-26
;  PRIOR APPLICATION NUMBER:  JP 1999-179043
;  PRIOR FILING DATE:  1999-06-24
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;  LENGTH: 2280
;  TYPE: DNA
;  ORGANISM: Arabidopsis thaliana
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (310)..(2142)
;  NAME/KEY: misc_feature
;  LOCATION: (1650)..(1655)
;  OTHER INFORMATION: BamHI recognition site
;  NAME/KEY: misc_feature
;  LOCATION: (1984)..(1989)
;  OTHER INFORMATION: SphI recognition site
US-09-339-947A-2

Alignment Scores:
Pred. No.:              2.23e-82      Length:          2280
Score:                  923.00        Matches:         216
Percent Similarity:     43.23%        Conservative:    52
Best Local Similarity:  34.84%        Mismatches:      108
Query Match:            38.81%        Indels:          244
DB:                     17            Gaps:            9

US-09-890-220-2 (1-445) x US-09-339-947A-2 (1-2280)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||   ::::::   |||   :::   |||   ||||||   |||:::   :::|||
Db        310 ATGTGCCATGAAGACTCCCGTCTGCGTATTTCGGAAGAGGAGGAGATTGCTGCTGAAGAG 369
```





RESULT 14
US-60-303-460-9
; Sequence 9, Application US/60303460
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Butler, Karlene
; APPLICANT: Niu, Xiping
; TITLE OF INVENTION: Flowering Regulating Polynucleotides and Polypeptides
; FILE REFERENCE: BB1493 US PRV
; CURRENT APPLICATION NUMBER: US/60/303,460
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Helianthus sp.
US-60-303-460-9

Alignment Scores:

Pred. No.:	8.68e-81	Length:	2346
Score:	907.50	Matches:	218
Percent Similarity:	43.19%	Conservative:	64
Best Local Similarity:	33.38%	Mismatches:	120
Query Match:	38.16%	Indels:	251

```
DB:                74            Gaps:          10

US-09-890-220-2 (1-445) x US-60-303-460-9 (1-2346)

Qy           1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
               |||||||||   :::   ||||||      :::   ||||||      :::   :::|||
Db         140 ATGTGCCGCAATGATGCTCGTGCTCATTTGACTCAGGAAGAACAAGCTGCAGCTGAGGAG 199

Qy          21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
               :::|||   :::|||||||||||||||   ||||||||||          |||::::::
Db         200 AGTCTTTCGGTTTACTGCAAACCTGTTGAATTATACAACATTCTCCAACGGCGCGCTGTT 259

Qy          41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                  ||||||   ||||||   ||||||||||:::||||||:::   ||||||:::|||:::
Db         260 AGAAATCCATTATTTCTTCAAAGATGTTTGCATTACAAGTTACAAGCGAAGCAGAAAAAA 319

Qy          61 Lys------------------------------------------------------- 61
               :::
Db         320 AGGATACAAATATCAGTATCTGTATCTGGGGCTACTAGTGATGGACATCAAACTCAGACG 379

Qy          62 ----------------------------------------------------------SerArg 63
                                                                       ::::::
Db         380 CTTTTTCCTTTGTACGTGTTATTGGCAAGACCAGTGTCTACTACAAATGTGGAGACACAG 439

Qy          64 SerThr------------------------------------------------------ 65
               ||||||
Db         440 AGTACAACTGTTTATCGTTTCAATCGAGCACGTAAGTTGACGGTTTTTGGTGGGGCTGAG 499

Qy          65 ---------------------------------------------------------- 65

Db         500 ACATCGAGTACAGCTCGAGCAAAGTTTATGCTTCCCGAGATGAATAAACTATCCAAAGAG 559

Qy          65 ---------------------------------------------------------- 65

Db         560 ATCAAATCTGGCTCTCTAGATGTATTGTTGGTTAGCTGTGCTGACTCCACAGATTCTCCG 619

Qy          65 ---------------------------------------------------------- 65

Db         620 AAGATTGACCTAACGGAGGATCATATGTTTTGTAACTCATTGAATTGTGAGGGTTACTGC 679

Qy          65 ---------------------------------------------------------- 65

Db         680 TTGTTTGGGAAGATTCCAATGGATTTACTTCATTCTTCATGGGAAAAATCTCCGACATTA 739

Qy          65 ---------------------------------------------------------- 65

Db         740 AGTTTAGGGGGAAGAGCTGAAATGATGTCAACTGTTATTATGCAGTCCTGCAGTATGAAG 799

Qy          65 ---------------------------------------------------------- 65

Db         800 TTGAGTTCTTCAAATGGGGAAAAGTGTGTATCTTTTCGGTTTCCATATAACTCTGAAGCT 859

Qy          65 ---------------------------------------------------------- 65

Db         860 GTGAGCATATTGCAGCAAGTACCGGTCGTCATTTCTGCTGAAGAATTTGGGGCAAAAAAT 919

Qy          65 ---------------------------------------------------------- 65

Db         920 ATGTTGCCCTATGATGTATATTCATACAACGATGCTCCTAGACCCGGTGTAATGAGATTG 979

Qy          66 ------GlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu 83
                     |||   ||||||||||||||||||      ||||||||||||   :::||||||
Db         980 CGGTCTGGAAATGTCGTTTTTAATTACAAGTACTACAACAATACATTGCATAGGACTGAA 1039

Qy          84 ValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeu 103
               |||   ||||||   ||||||||||||||||   :::   |||   |||:::|||||||||
Db        1040 GTTACAGAGGATTTTTCTTGTCCGTTCTGTTTGGTGAAATGCGCTAGTTACAAGGGCCTG 1099

Qy         104 GlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGlu 123
               :::|||||||||   ||||||||||||||||   :::||||||   :::   |||:::
Db        1100 AGATTTCACTTGACTTCATCACATGATCTCTTTCATTATGAGTTTTGGGTTACTGAAGAC 1159

Qy         124 TyrGlnThrValAsnValSerValLysLeu---------AsnSerPheIlePheGluGlu 140
                     |||   |||   ||||||::::::            :::|||   |||      |||
Db        1160 TGTCAGGTCGTGATTGTATCTATCAGAACCGATGCAAGCAGCTCTGAGATAATTCCTGAG 1219

Qy         141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
                     :::   |||   |||   :::   ||||||   |||   |||   |||:::|||
Db        1220 ------AACGTTGATCCGAAACAACAAATGTTCTCTTATTGCCATAAATCGAGAAGACGC 1273

Qy         161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLys---ValCysPheLeuProLeu 179
               |||:::         :::      |||:::             |||      |||   |||
Db        1274 CGGAAA------CCAAAAACTCAAACTCAAAACGCAAATCATGTGCATCCACTCGTGCTG 1327

Qy         180 AspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsn 199
               ||||||         |||                 ::::::   |||   |||
Db        1328 GATTCAACCATGTCTACA--------------GCTCTCAGTGAGCTTATAGATAATACA 1372

Qy         200 ArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsn 219
               |||                         :::|||         |||:::   :::
Db        1373 GATGGT----------------------GTAGCTGAGTGCATGGAGCTGGATGCATGT 1408

Qy         220 IleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIle--------- 236
               |||   |||   |||             :::|||      :::   |||
Db        1409 AGTCCGGATGCAAGTGCCACATGCCCCTCAAATGCCGAACCCGAGCCTATACAGTCATTA 1468

Qy         237 --------------LeuThrSerGluAlaValValProAlaThrLysThrArgLysLeu 251
                           |||         |||::::::         ||||||||||||||
Db        1469 CCTGAAAGCAGCAACCTTGCACCTCCAGCCATGCTTCAGTTTGCAAAGACGAGAAAGTTA 1528

Qy         252 SerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHis 271
               |||   ||||||||||:::   |||:::      ||||||||||||||||||||:::|||
Db        1529 TCGGTTGAGCGATCCGACCCTAGAAACCGAGCCCTGCTGCAGAAACGGCAGTTCTTTCAC 1588

Qy         272 SerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAsp 291
               |||||||||   ||||||||||||:::|||||||||   ::::::||||||||||||||||
Db        1589 TCCCACCGAGCTCAGCCAATGGCGATAGAACAAGTATTCGCTGAGCGAGACAGTGAAGAT 1648

Qy         292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAsp 311
               ||||||||||||||||||||||||   |||||||||:::||||||||||||||||||||||
Db        1649 GAAGTTGACGATGATGTTGCTGATCTCGAGGACCGTAGGATGCTGGACGACTTCGTGGAT 1708

Qy         312 ValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArg 331
               |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        1709 GTTTCCAAAGACGAGAAACAGATGATGCATCTATGGAACTCATTTGTGAGAAAGCAACGG 1768

Qy         332 ValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLys 351
               |||:::|||||||||||||||||||||:::|||||||||||||||||:::   :::   :::
Db        1769 GTGCTGGCGGATGGCCATATTTCGTGGTCGTGCGAAGCGTTTTCAAAGCTGCATGGTCAA 1828

Qy         352 GluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrp 371
               :::|||   :::      ||||||   ||||||||||||||||||:::||||||||||||
Db        1829 GACCTTGTTCAGGCTCCATCACTACTCTGGTGCTGGAGATTGTTTATGATAAAATTATGG 1888

Qy         372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
               ||||||||||||:::|||||||   ::::::||||||||||||   |||||||||
Db        1889 AATCATGGGCTCCTAGACTCGCGGAGCTTAAACAATTGTAATGTGATCCTTGAACAATAC 1948

Qy         392 Arg---------------AsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerVal 406
               :::                                                 :::   |||
Db        1949 CAGGGTCAGGTTCAGGGTCAAGGTCAAAGTCAAGTTGAAGGTCAAGGTCAAAGTCAAGTT 2008

Qy         407 AspArgProSerAspSerAsnThrAsnAsnAsnAsnAsnIle 419
               :::            |||   :::   |||   :::|||
Db        2009 GAAGGTGAAGGCCAAAGTCAAAGTCAAAATCAAGACATA 2047

RESULT 15
US-09-339-947A-9
; Sequence 9, Application US/09339947A
; GENERAL INFORMATION:
;  APPLICANT:  YOSHIDA, Nobumasa
;  APPLICANT:  YANAI, Yukihiro
;  APPLICANT:  KATO, Yoshihiro
;  APPLICANT:  HIRATSUKA, Junzo
;  APPLICANT:  TAKAHASHI, Shigeru
;  APPLICANT:  MIWA, Tatsushi
```

```
;  TITLE OF INVENTION:  GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
;  TITLE OF INVENTION:  FLOWERING
;  FILE REFERENCE:  032735-001
;  CURRENT APPLICATION NUMBER: US/09/339,947A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:  JP 1998-180065
;  PRIOR FILING DATE:  1998-06-26
;  PRIOR APPLICATION NUMBER:  JP 1999-179043
;  PRIOR FILING DATE:  1999-06-24
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 2248
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (86)..(1897)
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(41)
;   OTHER INFORMATION: NheI recognition site
US-09-339-947A-9

Alignment Scores:
Pred. No.:                1.03e-80      Length:          2248
Score:                    906.50        Matches:         209
Percent Similarity:       42.68%        Conservative:    59
Best Local Similarity:    33.28%        Mismatches:      109
Query Match:              38.12%        Indels:          251
DB:                       17            Gaps:            7

US-09-890-220-2 (1-445) x US-09-339-947A-9 (1-2248)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              |||||||||         ||||||:::   ||||||:::|||   ::::::   :::|||
Db         86 ATGTGCCGCCACCAGCCAAGGGCTCGGCTCTCTCCCGATGAGCAGCTTGCAGCTGAAGAA 145

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::       :::||||||||||||||   ||||||||||||         |||||||:::
Db        146 AGCTTCGCATTATACTGCAAGCCGGTCGAGTTGTATAATATCATTCAGCGCCGATCCATT 205

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||:::||||||   |||||||||   |||||||||   |||:::|||||||:::
Db        206 AAAAATCCTGCTTTTCTTCAAAGATGCCTTCTTTACAAGATTCACGCAAGACGGAAGAAG 265

Qy         61 LysSerArgSerThr---------------------------------------------- 65
              :::|||      |||
Db        266 AGGAGCCTGATAACCATATCACTTTCTGGAGGCACAAATAAAGAACTGCGGGCACAAAAT 325

Qy         65 ------------------------------------------------------------ 65

Db        326 ATCTTTCCTCTTTATGTTCTGTTAGCTAGACCTACTAATAATGTTTCACTTGAAGGGCAT 385

Qy         65 ------------------------------------------------------------ 65

Db        386 TCTCCGATATATCGATTCAGTCGTGCTTGTTTGTTGACTTCTTTTCATGAATTTGGAAAT 445

Qy         65 ------------------------------------------------------------ 65

Db        446 AAAGACTACACTGAAGCAACATTCGTCATTCCTGATGTGAAGAACTTAGCAACCTCCCGA 505

Qy         65 ------------------------------------------------------------ 65

Db        506 GCTTGCAGCCTTAATATTATCCTTATCAGCTGTGGACGAGCTGAGCAAACTTTTGATGAC 565

Qy         65 ------------------------------------------------------------ 65

Db        566 AATAACTGTTCTGGGAACCATGTGGAAGGCTCTACTCTCCAAAAGCTTGAAGGGAAGTGT 625

Qy         65 ------------------------------------------------------------ 65

Db        626 TTCTGGGGTAAAATACCAATCGATCTTCTTGCTTCATCTTTGGGAAATTGTGTGAGCTTA 685

Qy         65 ------------------------------------------------------------ 65

Db        686 AGTTTGGGACATACCGTGGAAATGTCTTCCACGGTTGAGATGACCCCAAGCTTCTTAGAG 745

Qy         65 ------------------------------------------------------------ 65

Db        746 CCAAAATTTCTGGAGGATGACAGTTGCTTGACATTTTGCTCTCAGAAGGTTGATGCTACT 805

Qy         65 ------------------------------------------------------------ 65

Db        806 GGTTCATTTCAACTGCAAGTTAGCATATCTGCTCAAGAGGCTGGTGCAAAAGACATGTCC 865

Qy         65 ------------------------------------------------------------ 65

Db        866 GAGTCTCCTTATAGTGTTTATTCATATAATGATGTGCCACCTTCGTCATTGACACATATT 925

Qy         66 --------------GlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
                           |||   |||:::|||||||||||||      ||||||||||:::|||
Db        926 ATAAGGTTGAGATCTGGCAATGTGCTTTTTAACTACAAATACTACAATAATACTATGCAA 985

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   :::   ||||||||||||
Db        986 AAAACCGAAGTCACTGAAGATTTTTCTTGCCCATTTTGCTTGGTACCATGTGGCAGCTTT 1045

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||      |||||||||:::|||||||||||||||   :::||||||   :::
Db       1046 AAGGGTCTAGGATGTCACCTAAACGCATCGCATGACCTTTTCCATTATGAGTTTTGGATA 1105

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu--- 139
                 ||||||   |||   ||||||||||||:::|||   :::|||:::      |||
Db       1106 TCTGAAGAGTGCCAGGCTGTTAATGTTAGTCTGAAGACTGATTCTTGGAGAACAGAGCTT 1165

Qy        140 ---GluGluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArg 158
                    ||||||      |||   :::   :::   ||||||      |||:::   :::
Db       1166 TTGGCTGAGGGAGTT---GATCCAAGACATCAAACATTTTCGTACCGCTCAAGATTTAAG 1222

Qy        159 LysArg---ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
              ||||||   |||                :::      |||   :::           :::
Db       1223 AAGCGTAAAAGGGTGGAAATCTCAAGTGATAAAATTAGGCATGTACATCCACATATTGTG 1282

Qy        178 ProLeuAspSerPro---SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsn 196
                    ||||||           |||:::|||:::
Db       1283 GATTCAGGATCACCTGAAGATGCCCAGGCAGGATCTGAAGACGATTACGTGCAGAGGGAA 1342

Qy        197 AspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 216
              :::|||:::
Db       1343 AATGGTAGT---------------------------------------------------- 1351

Qy        217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal--- 235
                          ::::::||||||:::|||:::|||          :::
Db       1352 ------------------TCTGTAGCACACGCTTCTGTTGATCCTGCTAATTCATTACAC 1393

Qy        236 ------IleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAla 253
                    |||::::::      |||:::          ||||||||||||||||||
Db       1394 GGTAGCAATCTTTCAGCACCAACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTT 1453

Qy        254 GluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
              ||||||::::::   |||:::      ||||||||||||||||||||:::|||||||||
Db       1454 GAACGAGCTGATCCCAGAAATCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCAC 1513

Qy        274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
              |||   ||||||||||||      :::|||   ||||||||||||||||||||||||||||
Db       1514 AGGGCTCAACCAATGGCATGGAGCAAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTT 1573

Qy        294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
              |||||||||:::||||||||||||||||||:::||||||||||||||||||||||||||
Db       1574 GATGATGACATTGCTGATTTTGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACA 1633

Qy        314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
              ||||||||||||      ||||||:::||||||||||||||||||||||||||||:::
Db       1634 AAAGACGAGAAACTTATTATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTA 1693

Qy        334 AlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353
```

```
                ||||||||||||||||   ||||||||||||||||||||:::|||:::   :::||||||
Db       1694 GCGGATGGCCATATTCCCTGGGCATGCGAAGCATTCTCGCAGTTTCATGGACAAGAACTT 1753

Qy        354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373
                 :::      :::|||   ||||||||||||   |||:::::::|||||||||||||||||
Db       1754 GTACAAAATCCAGCTCTACTATGGTGTTGGAGGTTTTTTATGGTCAAACTCTGGAACCAC 1813

Qy        374 GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
                 |||:::|||:::      :::|||   ||||||||||||||||||         |||
Db       1814 AGTCTACTGGATGCGCGAGCCATGAATGCCTGCAACACAATTCTTGAAGGCTACCTGAAC 1873

Qy        394 SerSerAspThrThrThrThrAsn 401
                 |||                |||
Db       1874 GGAAGCTCGGATCCAAAGAAAAAT 1897


Search completed: June 20, 2003, 02:18:49
Job time : 3231 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:51:48 ; Search time 956 Seconds
(without alignments)
3236.070 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7816032 seqs, 3476047932 residues

Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144853_16556/app_query.fasta_1.583
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220_@CGN_1_1_844_@runat_13062003_144853_16556 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1166	49.0	1722	10	US-10-425-114-11348	Sequence 11348, A
2	1166	49.0	1838	10	US-10-424-599-111422	Sequence 111422,
3	833	35.0	1841	10	US-10-425-114-31883	Sequence 31883, A
4	653	27.5	2640	9	US-10-231-778-6	Sequence 6, Appli
5	648.5	27.3	2700	10	US-10-437-963-91042	Sequence 91042, A
c 6	617.5	26.0	1004	10	US-10-437-963-11609	Sequence 11609, A
c 7	575.5	24.2	1070	10	US-10-424-599-141143	Sequence 141143,
8	492	20.7	2439	10	US-10-425-114-30709	Sequence 30709, A
9	479	20.1	716	10	US-10-424-599-111421	Sequence 111421,
10	472.5	19.9	443	5	US-09-837-604B-13507	Sequence 13507, A
11	472.5	19.9	443	7	US-09-837-604A-13507	Sequence 13507, A
12	435	18.3	558	10	US-10-389-048-21629	Sequence 21629, A
13	348.5	14.7	554	5	US-09-531-113-5697	Sequence 5697, Ap
14	348.5	14.7	554	6	US-09-531-113-5697	Sequence 5697, Ap
c 15	292	12.3	531	5	US-09-837-604B-940	Sequence 940, App
c 16	292	12.3	531	7	US-09-837-604A-940	Sequence 940, App
c 17	271	11.4	322	10	US-10-424-599-40645	Sequence 40645, A
18	264.5	11.1	6458	9	US-10-231-778-7	Sequence 7, Appli
c 19	256.5	10.8	632	10	US-10-424-599-73456	Sequence 73456, A
20	255.5	10.7	453	5	US-09-837-604B-6414	Sequence 6414, Ap
21	255.5	10.7	453	7	US-09-837-604A-6414	Sequence 6414, Ap
22	255.5	10.7	453	10	US-10-437-963-97425	Sequence 97425, A
23	221	9.3	3235	10	US-10-170-235-25585	Sequence 25585, A
24	204.5	8.6	3900	9	US-10-144-771-9790	Sequence 9790, Ap
25	191.5	8.1	250	5	US-09-531-113-40606	Sequence 40606, A
26	191.5	8.1	250	6	US-09-531-113-40606	Sequence 40606, A
27	190	8.0	939	10	US-10-437-963-81637	Sequence 81637, A
28	184	7.7	274	10	US-10-389-048-16407	Sequence 16407, A
29	159	6.7	921	10	US-10-424-599-108137	Sequence 108137,
30	153	6.4	3209	10	US-10-170-235-21113	Sequence 21113, A
31	140.5	5.9	350	10	US-10-424-599-9668	Sequence 9668, Ap
32	138	5.8	477	5	US-09-837-604B-52491	Sequence 52491, A
33	138	5.8	477	7	US-09-837-604A-52491	Sequence 52491, A
34	132	5.6	2250	10	US-10-369-493-45620	Sequence 45620, A
35	128	5.4	405	10	US-10-437-963-13671	Sequence 13671, A
36	124	5.2	3931	1	PCT-US02-38360-1	Sequence 1, Appli
37	123.5	5.2	564	10	US-10-437-963-7175	Sequence 7175, Ap
38	122	5.1	441	10	US-10-437-963-98795	Sequence 98795, A
39	119	5.0	3282	1	PCT-US01-29288A-215	Sequence 215, App
40	119	5.0	3282	10	US-10-149-310-215	Sequence 215, App
41	118	5.0	782	10	US-10-437-963-33351	Sequence 33351, A
42	115.5	4.9	10640	9	US-10-304-095-5	Sequence 5, Appli
43	115	4.8	2718	10	US-10-369-493-45577	Sequence 45577, A
44	114.5	4.8	2637	9	US-10-282-122A-16387	Sequence 16387, A
45	114	4.8	2892	1	PCT-US01-29288A-255	Sequence 255, App

ALIGNMENTS

RESULT 1
US-10-425-114-11348
; Sequence 11348, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11348
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701055296_FLI
US-10-425-114-11348

Alignment Scores:

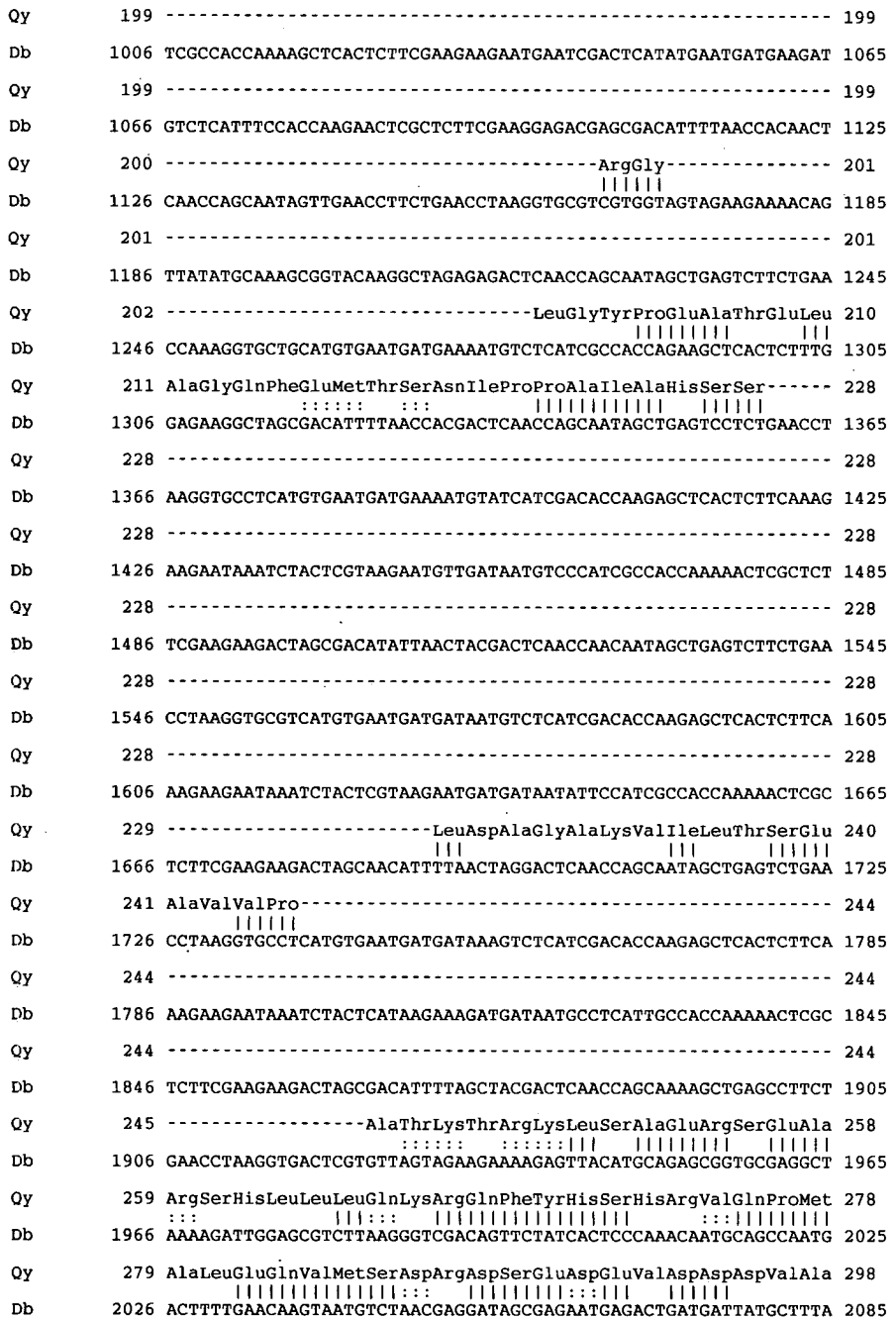
Pred. No.:	1.37e-116	Length:	1722
Score:	1166.00	Matches:	242
Percent Similarity:	65.65%	Conservative:	58
Best Local Similarity:	52.95%	Mismatches:	111

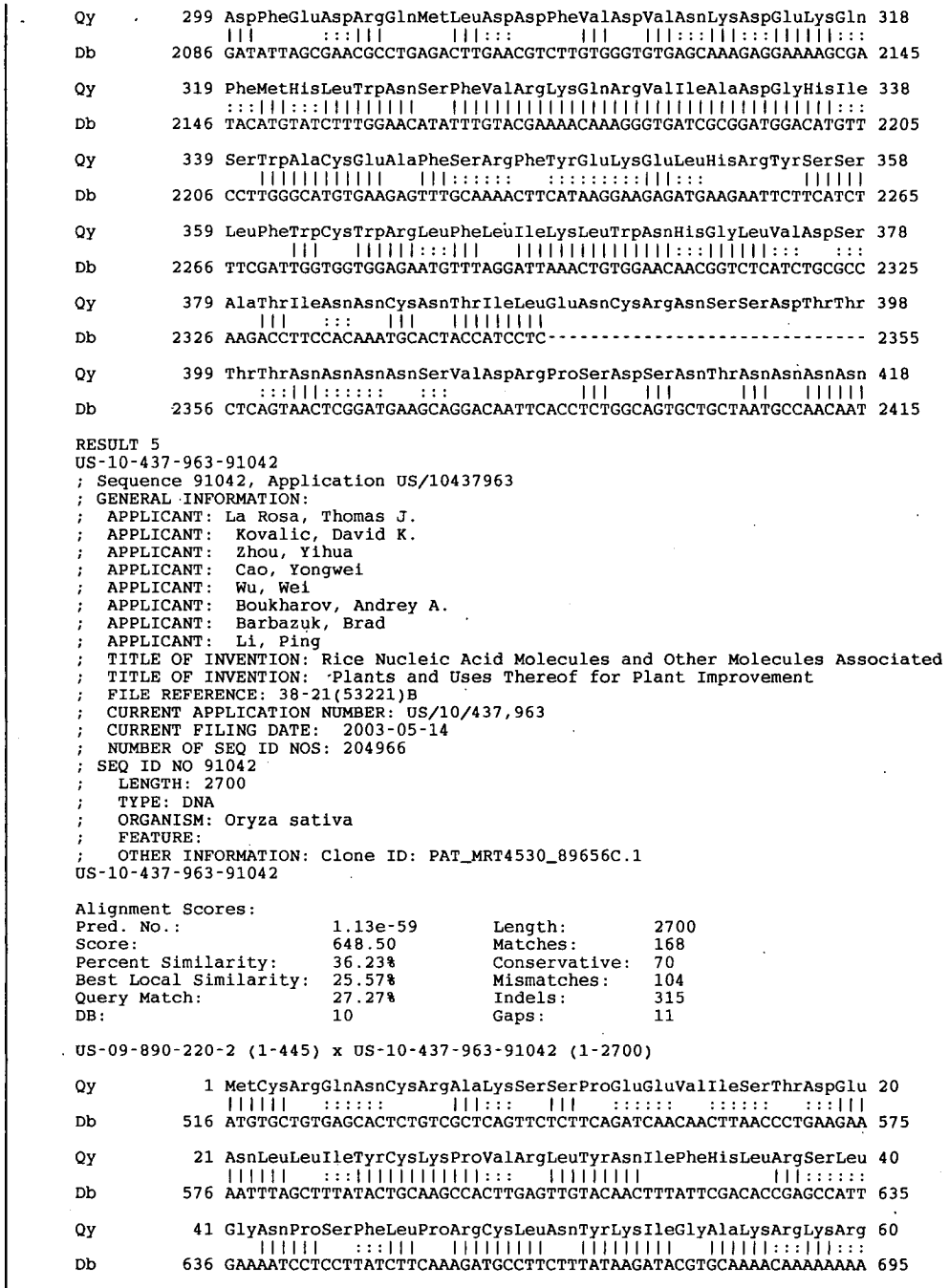
```
Query Match:            49.03%      Indels:        46
DB:                     10          Gaps:          9

US-09-890-220-2 (1-445) x US-10-425-114-11348 (1-1722)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||              :::   ||||||   |||:::   ||||||
Db        164 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 223

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        224 AGCCTTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCTCTT 283

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||   |||||||||
Db        284 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAAAGG 343

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              :::   |||:::   |||:::|||:::|||||||||:::|||      |||   |||:::
Db        344 AGGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTTCGA 400

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   |||   |||||||:::|||
Db        401 AAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAACTTT 460

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||:::|||||||||   |||||||||||||||||   |||||||||   :::
Db        461 AAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGGGTT 520

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
                 |||:::||||||   ||||||||||||||||||:::::   |||
Db        521 ACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGAT------ATATTGAGATCA 574

Qy        141 GluGlySerAspAsp---------AspLysPheGluProPheSerLeuCysSerLysPro 157
              |||         |||           :::   :::   |||      ||||||:::|||
Db        575 GAGAATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCT 634

Qy        158 ArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
              ||||||||||||::::::                 :::|||      |||   ||||||
Db        635 CGAAAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTG 694

Qy        178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197
                 ||||||||||||           |||      ||||||                :::|||
Db        695 GAGTTGGACTCACCA---------GAAGGCATACATAATGGATTTCTACAAAAAGATGAT 745

Qy        197 ------------------------------------------------------------ 197

Db        746 GATATCTTATCCTGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTC 805

Qy        198 --------------------GlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu 210
                                   |||      |||      :::|||         :::
Db        806 CCTAGTGGAAGAAATGACGGAGGAAAAATTTGGTCCTGATCATCCTGGCACCATGGACAAC 865

Qy        211 AlaGlyGlnPheGluMetThrSerAsnIleProPro---------AlaIleAlaHisSer 227
                    |||   :::   |||||||||           |||:::      |||
Db        866 CTGGAGCATGTGGAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCT 925

Qy        228 SerLeuAspAlaGly---AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246
              |||:::|||         :::|||   |||      |||:::      :::||||||
Db        926 TCTGTGGACCCTGAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCT--- 982

Qy        247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLys 266
              ||||||:::|||||||||   :::||||||::::::|||:::   :::||||||||||||
Db        983 AAAACAAAGAAGCTAAGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAG 1042

Qy        267 ArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAsp 286
              |||   |||:::|||||||||||||||||||||||||||||||||||||:::||||||
Db       1043 AGACTTTTCTTTCACTCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGAC 1102

Qy        287 ArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeu 306
              ||||||||||||||||||||||||||||||:::|||||||   |||||||||:::||||||
Db       1103 CGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTT 1162
```

```
Qy        307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
              ||||||||||||||||||||:::|||||||||||||||   ||||||||||||||||||||
Db       1163 GACGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTT 1222

Qy        327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346
              :::|||||||||||||||||:::|||||||||||:::   ||||||||||||||||||||
Db       1223 ATGAGGAAGCAAAGGGTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCC 1282

Qy        347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPhe 366
              :::   :::   |||||||||           :::|||||||||||||||||||||||||
Db       1283 AAGCTTCATGGAAAAGAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTC 1342

Qy        367 LeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
              :::|||||||||||||||||||||||||:::|||:::   |||:::|||||||||:::
Db       1343 ATGATCAAACTTTGGAATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATA 1402

Qy        387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerVal 406
              :::|||::::::   ||||||           |||     |||
Db       1403 GTATTAGATAGTTACAGAAATGAGGGATCGGGTACAAGAAAAAAT--------------- 1447

Qy        407 AspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisPro 423
                 ||||||      :::      ||||||   ||||||         |||
Db       1448 TGAAGACCATAAGGAAACTGACTCAACAACAAGAATATACAGTCTTCTCCG 1498

RESULT 2
US-10-424-599-111422
; Sequence 111422, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111422
;   LENGTH: 1838
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1
US-10-424-599-111422

Alignment Scores:
Pred. No.:              1.52e-116     Length:          1838
Score:                  1166.00       Matches:         242
Percent Similarity:     65.65%        Conservative:    58
Best Local Similarity:  52.95%        Mismatches:      111
Query Match:            49.03%        Indels:          46
DB:                     10            Gaps:            9

US-09-890-220-2 (1-445) x US-10-424-599-111422 (1-1838)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||              :::   ||||||   |||:::   ||||||
Db        258 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 317

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        318 AGCCTTTTGATTTATTGCAAGCCTGTTGAACTTTACAATATTCTCTACCGCCGTGCTCTT 377

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||   |||||||||
Db        378 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCGTTATAAAATAAGAGCAAGCCGTAAAAGG 437

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              :::   |||:::   |||:::|||:::|||||||||:::|||      |||   |||:::
Db        438 AGGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGACCGCTACAACATTCTTCGA 494

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
```

```
Db        495 AAGACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAACTTT 554

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120

Db        555 AAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGAGTTCTGGGTT 614

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140

Db        615 ACTGAAGATTACCAAGCAGTGAATGTCTCTGTAAAAATTGAT------ATATTGAGATCA 668

Qy        141 GluGlySerAspAsp---------AspLysPheGluProPheSerLeuCysSerLysPro 157

Db        669 GAGAATGTTGCTGATGGAGTAATTCCACAATCGCAAACCTTCTTCTTCTGTTCAAGACCT 728

Qy        158 ArgLysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177

Db        729 CGAAAACGTAGAAGAAAGGACTCTGTTCAAATTGAAAAACGCACCAATGTAAAATTCCTG 788

Qy        178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197

Db        789 GAGTTGGACTCACCA---------GAAGGCATACATAATGGATTTCTACAAAAAGATGAT 839

Qy        197 ------------------------------------------------------------ 197

Db        840 GATATCTTATCCTGCAAAGGAGAGAATGTGTCTAGAACATCTCGTAGTGAGAAGATTTTC 899

Qy        198 --------------------GlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu 210

Db        900 CCTAGTGGAAGAAATGACGGAGGAAAATTTGGTCCTGATCATCCTGGCACCATGGACAAC 959

Qy        211 AlaGlyGlnPheGluMetThrSerAsnIleProPro---------AlaIleAlaHisSer 227

Db        960 CTGGAGCATGTGGAATCCAGTTTCAACATTCCAGGTGTTTCAATTGCCATGCCCCAATCT 1019

Qy        228 SerLeuAspAlaGly---AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246

Db       1020 TCTGTGGACCCTGAATGTAGTAAATCAATATGTAAAAGTGATCCTGCTCTGCCTGCT--- 1076

Qy        247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLys 266

Db       1077 AAAACAAAGAAGCTAAGCATGGATCGATCAGACTCAAGAAACCGAATGCTTCTGCAGAAG 1136

Qy        267 ArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAsp 286

Db       1137 AGACTTTTCTTTCACTCACACAGAGTCCAGCCTATGGCACTAGAACAAGTGTTATCAGAC 1196

Qy        287 ArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeu 306

Db       1197 CGTGATAGTGAAGACGAAGTTGATGACGACATTGCAGATCTTGAAGATAGAAGGATGCTT 1256

Qy        307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326

Db       1257 GACGATTTTGTGGATGTTTCCAAAGATGAAAAACAGCTCATGCATCTCTGGAACTCTTTT 1316

Qy        327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346

Db       1317 ATGAGGAAGCAAAGGGTGCTAGCAGATGGTCATGTTCCGTGGGCCTGTGAGGCATTTTCC 1376

Qy        347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPhe 366

Db       1377 AAGCTTCATGGAAAAGAGCTGATCTCATCTCCAGCTTTATTTTGGTGTTGGAGGTTATTC 1436

Qy        367 LeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386

Db       1437 ATGATCAAACTTTGGAATCATGGTCTTCTTGATGCCTGTACAATGAACAACTGTAGCATA 1496

Qy        387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnAsnSerVal 406

Db       1497 GTATTAGATAGTTACAGAAATGAGGGATCGGGTACAAGAAAAAAT--------------- 1541

Qy        407 AspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisPro 423

Db       1542 TGAAGACCATAAGGAAACTGACTCAACAACAAGAATATACAGTCTTCTCCG 1592

RESULT 3
US-10-425-114-31883
; Sequence 31883, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31883
;   LENGTH: 1841
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12_FLI
US-10-425-114-31883

Alignment Scores:
Pred. No.:              4.12e-80        Length:         1841
Score:                  833.00          Matches:        180
Percent Similarity:     60.81%          Conservative:   45
Best Local Similarity:  48.65%          Mismatches:     83
Query Match:            35.03%          Indels:         62
DB:                     10              Gaps:           8

US-09-890-220-2 (1-445) x US-10-425-114-31883 (1-1841)

Qy         43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer 62

Db        459 CCTTCTCTATTACCAGAAATCATAAGGTTGAGA-------------------------- 491

Qy         63 ArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThr 82

Db        492 ------GCTGGAAATGTACTTTTTAATTACAAGTACTACAACAATACAATGCAAAAGACT 545

Qy         83 GluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGly 102

Db        546 GAAGTTACTGAAGATTTTTCCTGCCCATTTTGCTTGGTACCATGTGGAAGCTTCAAGGGC 605

Qy        103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 122

Db        606 CTAGGATGCCACTTAAATTCATCACATGACCTGTTCCACTATGAGTTGTGGATATCTGAA 665

Qy        123 GluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe------IlePheGluGlu 140

Db        666 GAGTGCCAGGCTGTGAATGTTAGTCTGAAGGCTGATGCCTGGAAAACTGAGTTTGTGGCA 725

Qy        141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160

Db        726 GAGGGAGTT---GATCCAAGGCATCAAACATTTTCCTACTGCTCGAGGTTTAAGAAGCGT 782

Qy        161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180

Db        783 AGA-----------AGGTTGGAAACCACAGCTGAGAAA------TTCAGGCATGTTCAC 824

Qy        181 SerProSerLeuThrAsnGlyThr---------------------------------- 188

Db        825 CCACATATTATGGAATCAGGATCACATGAAGATGCCCAGGCAGGGTCTGAAGATGACTAT 884

Qy        189 ---------GluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrPro 205

Db        885 GTGCAAAGGGAAAATGGGCTTTCTGTAGCAAAT-------------------------- 917

Qy        206 GluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAla 225
```

```
Db        918 ------------------------------ACTTCAGTTGACCCTAGTCAGCCTGTA 944

Qy        226 HisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValValProAla 245
              |||   |||                      |||:::         |||:::
Db        945 CATGGGAGC--------------------AATCTTTCACCACCAACAGTGCTACAGTTT 983

Qy        246 ThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuGln 265
              |||:::|||||||||||||||||||||||:::   |||:::     |||||||||
Db        984 GGAAAGTCAAGAAAACTATCTGCCGAGCGATCTGACCCCAGAAACCGACAACTGTTGCAA 1043

Qy        266 LysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSer 285
              ||||||||||||:::|||||||||||   |||||||||   ||||||||||||:::|||
Db       1044 AAGCGTCAGTTCTTCCATTCACACAGGGCACAGCCAATGCCATTGGAACAAGTCCTCTCA 1103

Qy        286 AspArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMet 305
              |||   |||||||||||||||||||||||||||:::|||||||||||||||||:::|||
Db       1104 GATCATGATAGCGAAGACGAGGTTGATGATGACATTGCCGACTTTGAGGATAGACGGATG 1163

Qy        306 LeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSer 325
              ||||||||||||||||||||   ||||||||||||       ||||||:::|||||||||
Db       1164 CTCGACGATTTTGTTGATGTTACAAAAGATGAGAAGCTTATTATGCATATGTGGAATTCA 1223

Qy        326 PheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPhe 345
              ||||||||||||||||||||:::|||||||||||||||   ||||||||||||||||||
Db       1224 TTTGTTCGAAAGCAAAGGGTGCTAGCCGATGGTCATATACCGTGGGCATGCGAGGCGTTC 1283

Qy        346 SerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeu 365
              ||||||   :::   :::::|||   :::      :::|||   |||   ||||||
Db       1284 TCCAGATGCCATGGACAACAGCTTGTACAAAACCCTGCCCTGCTGTGGGGCTGGCGTTTC 1343

Qy        366 PheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsn 385
              ||||:::|||||||||||||||||||   |||:::|||:::      :::|||   ||||||
Db       1344 TTCATGATTAAACTCTGGAACCACAGCCTTCTTGACGCCCGCGCCATGAATGCTTGCAAC 1403

Qy        386 ThrIleLeuGluAsnCysArgAsnSerSer 395
              |||||||||:::      ::::::   |||
Db       1404 ACAATCCTTCAAGGCTACCAAGACGGAAGT 1433

RESULT 4
US-10-231-778-6
; Sequence 6, Application US/10231778
; GENERAL INFORMATION:
;  APPLICANT: Bilodeau, Pierre
;  APPLICANT:  Chaudhury, Abdul M.
;  APPLICANT:  Dennis, Elizabeth S.
;  APPLICANT:  Koltunow, Anna M.G.
;  APPLICANT:  Luo, Ming
;  APPLICANT:  Peacock, William J.
;  TITLE OF INVENTION: Method for inducing seed development by down-regulating
;  TITLE OF INVENTION:  expression of the FIS2 gene
;  FILE REFERENCE: 72-98A
;  CURRENT APPLICATION NUMBER: US/10/231,778
;  CURRENT FILING DATE:  2002-11-08
;  PRIOR APPLICATION NUMBER: 09/398,237
;  PRIOR FILING DATE: 1999-09-20
;  PRIOR APPLICATION NUMBER: 60/101,184
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: AU PP6061
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6062
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PP6063
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: AU PQ1345
;  PRIOR FILING DATE: 1999-07-01
;  PRIOR APPLICATION NUMBER: AU PQ1346
;  PRIOR FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 239
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 2640
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2439)
;   OTHER INFORMATION: Nucleotides from 1 to 2439 represent protein
;   OTHER INFORMATION: coding sequence.
US-10-231-778-6

Alignment Scores:
Pred. No.:              3.53e-60        Length:         2640
Score:                  653.00          Matches:        184
Percent Similarity:     32.76%          Conservative:   65
Best Local Similarity:  24.21%          Mismatches:     147
Query Match:            27.46%          Indels:         364
DB:                     9               Gaps:           9

US-09-890-220-2 (1-445) x US-10-231-778-6 (1-2640)

Qy         13 GluGluValIleSerThrAspGluAsnLeuLeuIleTyrCysLysProValArgLeuTyr 32
              ::::::   :::     |||||||||:::::   |||   |||||||||   :::|||
Db        166 GATGATGATGTTGATGTTGATGAAAACATCATCAAATATATCAAACCTGTTGCAGTATAC 225

Qy         33 AsnIlePheHisLeuArgSerLeuGlyAsnProSerPheLeuProArgCysLeuAsnTyr 52
                          ||||||     ||||||  ||||||  |||   |||   |||
Db        226 AAGAAACTTGAAACTCGCTCAAAAAACAACCCATATTTCCTACGAAGGTCTTTGAAGTAC 285

Qy         53 LysIleGlyAlaLysArgLysArgLysSerArgSerThrGlyMetValValPheAsnTyr 72
              |||   ||||||:::|||:::||||||   |||   |||   :::   |||||||||
Db        286 ATAATCCAAGCAAAGAAAAAAAAGAAGTCAAATTCAGGTGGGAAAATAAGATTCAACTAC 345

Qy         73 LysAspCysAsnAsnThrLeuGln---LysThrGluValArgGluAspCysSerCysPro 91
              :::|||   :::|||   :::       |||   ||||||   |||:::   |||||||||
Db        346 AGGGATGTGAGTAACAAAATGACACTAAAAGCTGAAGTAGTGGAAAATTTTTCTTGCCCA 405

Qy         92 PheCysSerMetLeuCysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHis 111
              ||||||   :::   ||||||      :::|||||||||   ||||||   |||||||||
Db        406 TTTTGCTTGATTCCATGTGGAGGTCACGAGGGCTTGCAACTTCATTTGAAGTCATCACAT 465

Qy        112 AspLeuPheGluPheGluPheLysLeuPheGluGluTyrGln------ThrValAsnVal 129
              |||   |||:::|||||||||         |||:::                |||:::|||
Db        466 GACGCCTTTAAATTTGAGTTTTATCGGGCAGAGAAAGATCACGGACCGGAAGTTGATGTC 525

Qy        130 SerValLysLeuAsnSerPheIlePheGlu------GluGluGlySerAspAspAspLys 147
              |||||||||   ::::::      |||         ::::::         :::   :::
Db        526 TCCGTGAAAAGTGATACAATAAAATTTGGGGTTCTAAAGGATGATGTAGGAAATCCCCAA 585

Qy        148 PheGluProPheSerLeuCysSerLysProArg---LysArgArgGlnArgGlyGlyArg 166
                   |||   :::   |||||||||   |||   :::||||||||||||
Db        586 TTGAGCCCTTTGACGTTTTGCTCGAAAAATCGTAACCAAAGAAGACAAAGAGATGATAGC 645

Qy        167 AsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsn 186
              ||||||   ::::::|||   |||      :::   ||||||         |||
Db        646 AATAACGTTAAGAAACTTCCTGTACTCCTTATGGAGTTGGATTTAGATGACTTACCTCGT 705

Qy        187 GlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsn---------------------- 199
              ||||||||||||      |||   :::|||||||   |||
Db        706 GGAACAGAAAATGATTCTACTCATGTGAATGATGATAATGTCTCATCGCCACCAAGAGCT 765

Qy        199 ------------------------------------------------------------ 199

Db        766 CACTCTTCCGAGAAGATTAGCGACATTTTAACCACGACTCAACTAGCAATAGCTGAATCC 825

Qy        199 ------------------------------------------------------------ 199

Db        826 TCTGAACCTAAGGTGCCTCATGTGAATGATGGTAATGTCTCATCGCCACCAAGAGCTCAC 885

Qy        199 ------------------------------------------------------------ 199

Db        886 TCTTCGGCCGAGAAGAATGAATCTACTCATGTGAATGATGATGATGATGTCTCATCACCA 945

Qy        199 ------------------------------------------------------------ 199

Db        946 CCTAGAGCTCACTCTTTGGAGAAGAATGAATCTACTCATGTGAATGAGGATAATATTTCA 1005
```





RESULT 5
US-10-437-963-91042
; Sequence 91042, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91042
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_89656C.1
US-10-437-963-91042

Alignment Scores:

Pred. No.:	1.13e-59	Length:	2700
Score:	648.50	Matches:	168
Percent Similarity:	36.23%	Conservative:	70
Best Local Similarity:	25.57%	Mismatches:	104
Query Match:	27.27%	Indels:	315
DB:	10	Gaps:	11

US-09-890-220-2 (1-445) x US-10-437-963-91042 (1-2700)

```
Db        696 AGGATACAGATAACTATATCATTACCTGGAAGTAACAATAAGGAATTGCAAGCACAGAAT 755
Qy         61 ------------------------------------------------------------ 61
Db        756 ATCTTTCCTCTGTATGTTCTGTTTGCTAGACCTACTTCAAATGTTCCTATAGAAGGACAT 815
Qy         61 ------------------------------------------------------------ 61
Db        816 TCTCCAATATATCGGTTCAGTCAGGCCCGTTTGCTTACTTCCTTTAATGACTCTGGGAAT 875
Qy         61 ------------------------------------------------------------ 61
Db        876 AATGACCGTGCTGAAGCCACATTTGTTATTCCTGATCTGGAGACTTTAATTGCCACCCAA 935
Qy         62 --------------------------SerArgSerThr---------------------- 65
Db        936 GCTTATGGTCTTACTTTTATCCTTGTTAGCCGCGGTACCAAAAAAAATAAAGGGCGAACT 995
Qy         65 ------------------------------------------------------------ 65
Db        996 GGACAAAATCTTTGTGAAAATGACTGTTCTGAGAAACATGTGGACTACTCTTCTCTCCGA 1055
Qy         65 ------------------------------------------------------------ 65
Db       1056 AAGCTTGCAGGGAAATGTTTCTGGGGTAAAATTCCGATCACTTTACTTAATTCATCTTTG 1115
Qy         65 ------------------------------------------------------------ 65
Db       1116 GAGACTTGTGCGGATTTAATTTTGGGGCATATAGTGGAGTCACCTATCAGTATATGTATG 1175
Qy         65 ------------------------------------------------------------ 65
Db       1176 AGCCCAGGCTACTTAGAGCCAACATTTCTTGAGCATGACAATTGCTTGTCATTTTGTTCC 1235
Qy         65 ------------------------------------------------------------ 65
Db       1236 CGTAAAGCTGATGCTATGGTTCCATATCAGTTGCAAGTAAAAGTATCTGCAGCAGAGGCT 1295
Qy         65 ------------------------------------------------------------ 65
Db       1296 GGTGCAAAAGACATACTCAAATCTCCGTATAATTCCTTCTCATATAGTGACGTCCCACCA 1355
Qy         65 ------------------------------------------------------------ 65
Db       1356 TCCTTATTACTGCGTATTGTAAGGGCACTGCAGTACTATGCTAATACCATGAAGTTCTTT 1415
Qy         66 ------------------GlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeu 79
Db       1416 TTGTACAGGCTAAGAGTTGGAAATGTGCTCTTTAACTACAAG---------AACACA--- 1463
Qy         80 GlnLysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySer 99
Db       1464 CAAATGAGTGAAGTAACGGAAGATTTTACTTGCCCGTTTTGCTTGGTACGGTGTGGGAAC 1523
Qy        100 PheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLys 119
Db       1524 ------------------------------------------------------TTCAAG 1529
Qy        120 LeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu 139
Db       1530 ATATCTGAAGACTACCAGGCTGTTAATGTTACGCTGAAGAAAGATAACATGAGAACAGAG 1589
Qy        140 GluGluGlySerAsp---AspAspLysPheGluProPheSerLeuCysSerLysProArg 158
Db       1590 TTTGTGGCAGCAGAAGTTGATAATAGCCATCGGATCTTTTACTACCGATCAAGGTTTAAA 1649
Qy        159 LysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuPro 178
Db       1650 AAGAGTAGAACAGAA----------------------------------ATACTTCCG 1673
Qy        179 Leu---------AspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeu 195
```

```
Db       1674 GTTGCGCGTGCAGATGCACATATTATGGAATCAGGATCA---------------------- 1712
Qy         196 AsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGlu 215
Db       1713 --------------------------CCTGAAGAAACACAAGCGGAGTCTGAGGATGAT 1745
Qy         216 MetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal 235
Db       1746 GTCCAAGAGGAAAATGAAAACGCTTTGATTGATGATTCTAAGAAATTACATGGTAGCAAT 1805
Qy         236 IleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArg 255
Db       1806 CATTCACAATCAGAA---TTTCTGGCATTTGGGAAATCAAGGAAGCTATCAGCAAATCGA 1862
Qy         256 SerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgVal 275
Db       1863 GCTGATCCCAGAAATCGTCTACTTCTGCAAAAACGTCAGTTCATCCATTCTCATAAGGCA 1922
Qy         276 Gln------------------------------------------------------- 276
Db       1923 CAGACCCCTCTAGGAGATGACAAGGTGACCAACTTGGTGACGGCGGATGCCATGTCGGTA 1982
Qy         277 ProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
Db       1983 CCAATGACATTCGAAGAAGTTCTCTCAGATAATGATAGTGAAGATGAAGTAGATGATGAT 2042
Qy         297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
Db       2043 ATTGCTGATTTGGAAGATAGAAGGATGCTTGATGATTTTGTTGATGTTACAAAAGATGAG 2102
Qy         317 LysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
Db       2103 AAGCGCATTATGCACATGTGGAATTCATTTATTCGAAAACAAAGTATACTAGCTGATAGT 2162
Qy         337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
Db       2163 CACGTACCTTGGGCTTGTGAGGCATTCTCCCGACATCATGGAGAAGAACTTTTAGAAAAC 2222
Qy         357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
Db       2223 TCCGCTTTACTATGGGGATGGCGCATGTTTATGATCAAACTCTGGAATCACAGTCTACTC 2282
Qy         377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
Db       2283 TCTGCCCGCACAATGGACACCTGCAACAGAATTCTTGATGACATAAAAAAT 2333
```

RESULT 6
US-10-437-963-11609/c
; Sequence 11609, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated Wit
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 11609
; LENGTH: 1004
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609

Alignment Scores:

```
Pred. No.:              6.13e-57        Length:         1004
Score:                  617.50          Matches:        114
Percent Similarity:     76.92%          Conservative:   26
Best Local Similarity:  62.64%          Mismatches:     39
Query Match:            25.97%          Indels:         3
DB:                     10              Gaps:           1

US-09-890-220-2 (1-445) x US-10-437-963-11609 (1-1004)

Qy        223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal---------IleLeuThrSer 239
Db        844 TCTGTAGCACACGCTTCTGTTGATCCTGCTAATTCATTACACGGTAGCAATCTTTCAGCA 785

Qy        240 GluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
Db        784 CCAACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGA 725

Qy        260 SerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAla 279
Db        724 AATCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCA 665

Qy        280 LeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAsp 299
Db        664 TTGGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGAT 605

Qy        300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPhe 319
Db        604 TTTGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATT 545

Qy        320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
Db        544 ATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCC 485

Qy        340 TrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeu 359
Db        484 TGGGCATGCGAAGCATTCTCGCAGTTTCATGGACAAGAACTTGTACAAAATCCAGCTCTA 425

Qy        360 PheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAla 379
Db        424 CTATGGTGTTGGAGGTTTTTTATGGTCAAACTCTGGAACCACAGTCTACTGGATGCGCGA 365

Qy        380 ThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 399
Db        364 GCCATGAATGCCTGCAACACAATTCTTGAAGGCTACCTGAACGGAAGCTCGGATCCAAAG 305

Qy        400 ThrAsn 401
Db        304 AAAAAT 299

RESULT 7
US-10-424-599-141143/c
; Sequence 141143, Application US/10424599
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141143
;   LENGTH: 1070
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1
US-10-424-599-141143

Alignment Scores:
Pred. No.:              2.66e-52        Length:         1070
Score:                  575.50          Matches:        123
Percent Similarity:     63.14%          Conservative:   38
Best Local Similarity:  48.24%          Mismatches:     63
Query Match:            24.20%          Indels:         31
DB:                     10              Gaps:           6

US-09-890-220-2 (1-445) x US-10-424-599-141143 (1-1070)

Qy        145 AspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgArgGlnArgGly 164
Db       1032 GATCCAAGAGTGCAAACATTTTTCTTTTGTGGAAAGCCTCTAAAGCGTAGGACAACAGCA 973

Qy        165 GlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeu 184
Db        972 GACCAATCT-----------TTGAAAAATGCA---GTGGGCTTAGAGTCTTCCTTTCCT 928

Qy        185 ThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyr 204
Db        927 GCAGGAGGGACTGAT------ATTTTGGAGAAGGATGATGGT----------------- 892

Qy        205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIle 224
Db        891 -----------------------------------------ATTTCTGCCACAATT 877

Qy        225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThr------SerGluAlaVal 242
Db        876 ATTCGATCACGTCCTGATCGAGACTCTGTTCAGTCAATGTCTGACTGTGATCAAGCAGTG 817

Qy        243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeu 262
Db        816 CTTCAGTTTGCCAAGACAAGGAAGTTGTCAATTGAGCGTCCTGACCCACGAAACAGTACC 757

Qy        263 LeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeuGluGln 282
Db        756 TTCTTGAGGAAGCGACAATTTTTTCATTCACACAAAGCTCAGCCAATGGCAATTGAACAA 697

Qy        283 ValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAsp 302
Db        696 GTTCTATCCGATAAAGATAGCGAAGATGAAGTTGATGATGATGTTGCCGATTTTGAAGAT 637

Qy        303 ArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeu 322
Db        636 CGAAGGATGCTTGAAAATGTTGTTGATGTGAGCAATGATGAGAAGACTTTCATGCATATG 577

Qy        323 TrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrpAlaCys 342
Db        576 TGGAACTCATTTGTTCGGAAGCATCGTGTGATTGCAGATGGTCACATTTCATGGGCATGT 517

Qy        343 GluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCys 362
Db        516 GAGGCTTTCTCAAAATTGCATGCACCTGAGTTTGTTCAATCTCCCTCACTGGCAGGGTGT 457

Qy        363 TrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsn 382
Db        456 TGGAGAATATTTATGGTCAAATTATACAATCATGGTCTTCTAGATGCTCGGACCATGAAT 397

Qy        383 AsnCysAsnThrIleLeuGluAsnCys---ArgAsnSerSerAsp 396
Db        396 GACTGTAATATTATTCTTGAGCAATACCAAAGGCAGAATTCAGAT 352

RESULT 8
US-10-425-114-30709
; Sequence 30709, Application US/10425114
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
```

```
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30709
;   LENGTH: 2439
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73067D07_FLI
US-10-425-114-30709

Alignment Scores:
Pred. No.:                1.28e-42      Length:          2439
Score:                    492.00        Matches:         160
Percent Similarity:       29.74%        Conservative:    55
Best Local Similarity:    22.13%        Mismatches:      99
Query Match:              20.69%        Indels:          411
DB:                       10            Gaps:            12

US-09-890-220-2 (1-445) x US-10-425-114-30709 (1-2439)

Qy          3 ArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGluAsnLeu 22
              ||||||      |||         ||||||:::|||   ::::::   ::::::::::|||
Db        283 CGTCAGCAGCTAAGAACTGCATTGTCTCCAGATGAGAAACTTACCGCTGAAAAAGATTTG 342

Qy         23 LeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeuGlyAsn 42
              :::|||||||||||||||||   ||||||||||||         |||::::::   |||
Db        343 GCTCTGTATTGCAAGCCAGTCGAGCTCTACAATATTATTCAACGGCGAGCCATGAAAAAT 402

Qy         43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg------ 60
              |||    |||:::   |||||||||   |||   |||   |||:::|||||||:::
Db        403 CCCCTTTTTATTCAAAGATGCCTTCTTTATAATATACATGCGAGGAGGAAAAAGAGGATT 462

Qy         60 ------------------------------------------------------------ 60

Db        463 CAGATAACCATATCACTTTCTGGAAGTACAAATACTGAGTTGCAAACACATTATGTCTTT 522

Qy         60 ------------------------------------------------------------ 60

Db        523 CCTCTTTATGTTCTGTTAGCTAGACCCACTAGTAACCTTTCACTTGAAGGGCATTCTCCA 582

Qy         60 ------------------------------------------------------------ 60

Db        583 ATTTATCGATTCAGTCGGGTTTGCTTGCTTACTTCCTTTAGTGAACATGGAAATAAGGAC 642

Qy         61 ---------------------------------LysSerArgSerThr------------ 65
                                               ||||||   ||||||
Db        643 AACAGTGAAGCTACATTCATCATTCCTGACGTGAAGAGTTTGTCAACCTCCCGTGCTTGC 702

Qy         65 ------------------------------------------------------------ 65

Db        703 AACCATGATATTATCTTTATTAGCTGTGGGCAAGTTGGACAAAGTAATGGTGAAGATAAC 762

Qy         65 ------------------------------------------------------------ 65

Db        763 TGCTCTGGGAACCATGTGGAAGATTCTTCTCTCCAAATGCTTGAAGGGAAATGCTCCTGG 822

Qy         65 ------------------------------------------------------------ 65

Db        823 GGTAAAATACCAACTAATTTACTTGCTTCATCTTTGGAGAGTTGTGTCAATTTAAGTTTG 882

Qy         65 ------------------------------------------------------------ 65

Db        883 GGACATATTGTGGAGTTGGCATCTAAAGTTACAATGAGACCAAGCTTCTTAGAGCCAAAA 942

Qy         65 ------------------------------------------------------------ 65

Db        943 TTTCTGGAGCAAGACAGTTGCTTGACATTTTGCTCTCATAAGGTTGATGCTGTGGGTTCA 1002

Qy         65 ------------------------------------------------------------ 65

Db       1003 TATAAACTACAACTATGCATGTCCGCACAAGAGGCTGGTGCAAGAGATATGTCTTTGTCT 1062

Qy         65 ------------------------------------------------------------ 65

Db       1063 CCATATAGTAGTTACTCATATAATGATGTCCCACCTTCGTCATTATCAGATATCATAAGG 1122

Qy         66 ---------GlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThr 82
                      |||   |||:::|||||||||||      ||||||||||:::|||:::|||
Db       1123 TTAAGATCTGGCAATGTACTTTTTAATTACAAGTACTACAATAATACAATGCAAGAGACT 1182

Qy         83 GluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGly 102
              ||||||   ||||||   ||||||||||||||   :::   ||||||||||||||||||
Db       1183 GAAGTCACTGAAGATTTCTCTTGTCCATTTTGCTATGTACGATGTGGAAGCTTCAAGGGT 1242

Qy        103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGlu-PheLysLeuPhe-- 121
              |||      ||||||||||||||||||||||||||||   :::|||  |||   |||
Db       1243 CTAGGATGCCATTTAAACTCATCACATGATCTATTCCACTATGAGTTTTGGGTTGATATC 1302

Qy        122 -GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe-------------- 136
               ||||||||||||   ||||||||||||:::|||   :::::::::
Db       1303 TGAAGAGTACCAGGTTGTTAATGTTAGTCTGAAGGCTGATGCTTGGAGAACAGAGTGTTC 1362

Qy        137 ---------------------------------IlePheGluGluGluGly-------- 142
                                              :::   :::   ||||||
Db       1363 TTCCAAAGGTTCGTTTAAATGGAGGCTCCACTTAGTTGGGCAAACCGAAGGCATTGCATG 1422

Qy        143 ----------SerAspAspAsp---------LysPheGluProPheSerLeuCysSerLy 156
                       ||||||                 |||   ||||||      |||   ||
Db       1423 TTCCTTAGTCCAGGACGACTCCCTGCTTTGTGATTTCCTTCCTTTCTTGAG-TGCAG-AA 1480

Qy        156 sProArg------------------------------------------------------ 158
              |||||||
Db       1481 GCCAAGGATGAAAGAGCATTGGCTATTCTATATCTCCAATCAAATTCCTTACATAGTGTA 1540

Qy        158 ------------------------------------------------------------ 158

Db       1541 TTGTGTACCTCAATGATTTTTTGGTGGATAATGTTTTACATGCTTAACAATGCATGGTTA 1600

Qy        158 ------------------------------------------------------------ 158

Db       1601 CTATTTCTTTAAATGGAAGCACTCTTAACTGTATTTGGTAACTAATGATCCTAACTGGCA 1660

Qy        158 ------------------------------------------------------------ 158

Db       1661 GCTTTTTGCGGAGGGCGTTGATCCAAGGCATCAAACATTTTCTTATCGCTCAAGGTTTAA 1720

Qy        159 ----LysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLe 177
                  ::::||||||   :::                  |||   :::          ::
Db       1721 GAAGCGTAGACGATCAAAGAACACAATGGAGAAAATCAGGCATGTACACTCACATATTAT 1780

Qy        177 uProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAs 197
              :          ||||||            |||:::|||:::
Db       1781 GGAATCAGGTTCACCTGAAGATGAGGCAGGATCTGAGGACAACTTTGTG---------CA 1831

Qy        197 pGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetTh 217
               |||      |||
Db       1832 AGGGGAGAATGGG---------------------------------------------- 1844

Qy        217 rSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal----- 235
                          ::::::|||::::::|||:::|||          :::
Db       1845 ------------ACTTCTGTAGCAAATGCTTCGATTGATCCTGCTCAATCTTTACATGG 1891

Qy        236 ----IleLeuThrSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGl 254
                  |||:::         |||:::         ||||||||||||||||||   ||
Db       1892 CAGCAATCTTTCACCACCAACAGTACTACAGTTTGGGAAGACAAGGAAGCTATCT---GA 1948

Qy        254 uArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisAr 274
              |||||||:::   |||:::      ||||||||||||||||||||:::||||||||||
Db       1949 GAGATCTGACCCTAGAAATCGGCAACTCCTGCAAAAACGACAGTTCTTCCATTCTCACAG 2008

Qy        274 gValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAs 294
              |
Db       2009 G------------------------------------------------------------ 2009

Qy        294 pAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLy 314
```

```
Db       2009 ------------------------------------------------------------ 2009

Qy        314 sAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGln-ArgValIleA 334
                                                                 ||||||:::|
Db       2010 ----------------------------------------------GCGCAGAGTGTTAG 2023

Qy        334 laAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuH 354
              ||||||||||||||   |||||||||||||||||||||:::   :::   :::::::|||
Db       2024 CTGATGGTCATATACCTTGGGCCTGCGAGGCATTCTCCCAGTTGCATGGACGACAACTTA 2083

Qy        354 isArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisG 374
                :::      :::|||   |||   ||||||   |||:::|||||||||||||||||||
Db       2084 TACAAAATCCTGCTCTGCTGTGGGGTTGGCGTTTCTTCATGATTAAACTTTGGAACCATA 2143

Qy        374 lyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnS 394
                ::::::|||:::   |||:::|||   ||||||||||:::|||:::      :::
Db       2144 ACATTTTAGATGCCCGCACTATGAACACATGCAATACAGTCCTTCAAATTTTACAAGAAG 2203

Qy        394 erSer 395
                |||
Db       2204 AAAGC 2208
```

RESULT 9
US-10-424-599-111421
; Sequence 111421, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111421
; LENGTH: 716
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1
US-10-424-599-111421

Alignment Scores:

Pred. No.:	5.19e-42	Length:	716
Score:	479.00	Matches:	100
Percent Similarity:	72.46%	Conservative:	21
Best Local Similarity:	59.88%	Mismatches:	39
Query Match:	20.14%	Indels:	7
DB:	10	Gaps:	3

US-09-890-220-2 (1-445) x US-10-424-599-111421 (1-716)

```
Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||                :::   ||||||   |||:::   ||||||
Db        217 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 276

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        277 AGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCTCTT 336

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||:::||||||||||
Db        337 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAAAGG 396

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              :::   |||:::   |||:::|||:::|||||||||:::|||       |||   |||:::
Db        397 AGGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTTCGG 453

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   |||   ||||||||||||
```

```
Db        454 AAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGAAGCTTT 513

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
              |||||||||:::|||||||||   ||||||||||||||||||   |||   |||   :::
Db        514 AAGGGTTTGCGATTTCATCTTTGTTCATCACATGATCTATTCAACTTTGGGTTCTGGGTT 573

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu-Gl 140
              |||:::||||||   ||||||||||||||||||:::|||      |||||||||
Db        574 ACTGAAGATTACCAAGCAGTGAATGTCTCCGTGAAAATTAAT------ATATTTGAGATC 627

Qy        140 uGluGlySerAspAsp---------AspLysPheGluProPheSerLeuCysSerLysPr 157
              |||      |||          :::   :::   |||      ||||||:::||
Db        628 AGAGAATGTTGCTGATGGAGTAATTCCACAATCGCAAACTTTCTTCTTTTGTTCAAGACC 687

Qy        157 oArgLysArgArgGlnArg 163
              ||||   ||||||::::::
Db        688 TCGAAACCGTAGAAGAAAG 706
```

RESULT 10
US-09-837-604B-13507
; Sequence 13507, Application US/09837604B
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: De La Pena, Robert C.
; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604B
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604B-13507

Alignment Scores:

Pred. No.:	1.28e-41	Length:	443
Score:	472.50	Matches:	90
Percent Similarity:	80.60%	Conservative:	18
Best Local Similarity:	67.16%	Mismatches:	23
Query Match:	19.87%	Indels:	3
DB:	5	Gaps:	1

US-09-890-220-2 (1-445) x US-09-837-604B-13507 (1-443)

```
Qy        223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal---------IleLeuThrSer 239
              ::::::||||||:::|||:::|||            :::           |||::::::
Db         36 TCTGTAGCACACGCTTCTGTTGATCCTGCTAATTCATTACACGGTAGCAATCTTTCAGCA 95

Qy        240 GluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
                    |||:::         ||||||||||||||||||   ||||||::::::   |||
Db         96 CCAACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGA 155

Qy        260 SerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAla 279
              :::      ||||||||||||||||||||:::|||||||||||   ||||||||||||
Db        156 AATCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCA 215

Qy        280 LeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAsp 299
              ||||||||||||   ||||||||||||||||||||||||||||||||||:::||||||
Db        216 TTGGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGAT 275

Qy        300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPhe 319
              ||||||||||||:::||||||||||||||||||||||||   ||||||||||||
Db        276 TTTGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATT 335
```

```
Qy        320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
              ||||||:::||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        336 ATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCC 395

Qy        340 TrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353
              ||||||||||||||||||      |||:::   :::|||||||
Db        396 TGGGCATGCGAAGCATTTCTCGAGTTTCATGGACAAGAACTT 437

RESULT 11
US-09-837-604A-13507
; Sequence 13507, Application US/09837604A
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  APPLICANT:  De La Pena, Robert C.
;  APPLICANT:  Bougri, Olegs
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(51892)B
;  CURRENT APPLICATION NUMBER: US/09/837,604A
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: US 60/197,872
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
;   LENGTH: 443
;   TYPE: DNA
;   ORGANISM: Oryza sativa nipponbare
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604A-13507

Alignment Scores:
Pred. No.:                1.28e-41      Length:          443
Score:                    472.50        Matches:         90
Percent Similarity:       80.60%        Conservative:    18
Best Local Similarity:    67.16%        Mismatches:      23
Query Match:              19.87%        Indels:          3
DB:                       7             Gaps:            1

US-09-890-220-2 (1-445) x US-09-837-604A-13507 (1-443)

Qy        223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal---------IleLeuThrSer 239
              ::::::|||||||:::|||:::|||          :::          |||:::::
Db         36 TCTGTAGCACACGCTTCTGTTGATCCTGCTAATTCATTACACGGTAGCAATCTTTCAGCA 95

Qy        240 GluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
                 |||:::         ||||||||||||||||||  ||||||::::::   |||
Db         96 CCAACAGTGTTACAGTTTGGGAAGACAAGAAAGCTGTCTGTTGAACGAGCTGATCCCAGA 155

Qy        260 SerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAla 279
              :::      ||||||||||||||||||||:::||||||||||||   ||||||||||||
Db        156 AATCGGCAGCTCCTACAAAAACGCCAGTTCTTTCATTCTCACAGGGCTCAACCAATGGCA 215

Qy        280 LeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAsp 299
              ||||||||||||   ||||||||||||||||||||||||||||||||||:::||||||
Db        216 TTGGAGCAAGTTTTCTCAGATCGTGATAGTGAAGATGAAGTTGATGATGACATTGCTGAT 275

Qy        300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPhe 319
              ||||||||||||:::||||||||||||||||||||||||   ||||||||||||
Db        276 TTTGAAGATAGAAGAATGCTTGATGATTTTGTTGATGTTACAAAAGACGAGAAACTTATT 335

Qy        320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
              ||||||:::||||||||||||||||||||||||||||||:::|||||||||||||||
Db        336 ATGCATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCC 395

Qy        340 TrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeu 353
              ||||||||||||||||||      |||:::   :::|||||||
Db        396 TGGGCATGCGAAGCATTTCTCGAGTTTCATGGACAAGAACTT 437

RESULT 12
US-10-389-048-21629
```

```
; Sequence 21629, Application US/10389048
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka
;  APPLICANT:  Shenk, Michael A.
;  TITLE OF INVENTION: Polynucleotides, Materials Incorporating
;  TITLE OF INVENTION:  Them, and Methods for Using Them
;  FILE REFERENCE: 11000.1041U1c1
;  CURRENT APPLICATION NUMBER: US/10/389,048
;  CURRENT FILING DATE:  2003-03-13
;  NUMBER OF SEQ ID NOS: 25129
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21629
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-10-389-048-21629

Alignment Scores:
Pred. No.:                2.31e-37      Length:          558
Score:                    435.00        Matches:         93
Percent Similarity:       68.29%        Conservative:    19
Best Local Similarity:    56.71%        Mismatches:      46
Query Match:              18.29%        Indels:          8
DB:                       10            Gaps:            3

US-09-890-220-2 (1-445) x US-10-389-048-21629 (1-558)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||   |||:::               |||   |||:::      :::   :::|||
Db         74 ATGTGCCATCAAAGTTCTTGTGAACATTTTTCTGATGAAGATGCAACTGCAGCAGAAGAG 133

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              ||||||||||:::|||||||||||||||   ||||||||||||   |||   |||::::::
Db        134 AATCTCCTGGTGTATTGCAAGCCTGTTGAGCTATACAACATTCTACACCGCCGTGCCATA 193

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||||   |||||||||  :::|||||||  |||:::|||||||||
Db        194 CATAATCCTTCTTTCCTCCGGAGATGTTTGCGTTTCAAAATACAAGCAAGACGTAAGCGT 253

Qy         61 LysSerArg---SerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeu 79
                 |||         |||:::|||:::||||||||:::|||   ||||||
Db        254 AATAAGAGGTTGAGGGCTGGAATAGTCATTTTCAACTATAGGGACCGCAACAACGTGCAT 313

Qy         80 GlnLysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySer 99
              ||||||   ||||||   ||||||   ||||||||||||||   |||   |||   |||
Db        314 CAAAAGGTTGAAGTAACTGAAGATTTCTCTTGCCCATTTTGCTTAATGCAATGTGCAAGC 373

Qy        100 PheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLys 119
              |||||||||||:::   ||||||   ||||||||||||||||   |||||||||
Db        374 TTTAAGGGTCTAAGGTTGCATTTATGCTCATCGCATGACTTGTTCAACTTCGAGTTTTGG 433

Qy        120 LeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu 139
              :::   ||||||||||   ||||||||||||||||||:::::      :::   |||
Db        434 GTTATTGAAGAGTATCAGGCAGTGAATGTCTCTGTTAAAGTTGAT------ATGCT-GAG 486

Qy        140 GluGluGlySerAsp---------AspAspLysPheGluProPheSerLeuCysSerLys 156
                    |||||||||         :::   |||         |||      |||   |||
Db        487 ATCCGAGGTTCTGACTGAAGAAGCAACCCTAAATTGTGGACATTCTTCTTCTGCCC-AAA 545

Qy        157 ProArgLysArg 160
              ||||||:::|||
Db        546 CCACGAAGGCGG 557

RESULT 13
US-09-531-113-5697
; Sequence 5697, Application US/09531113
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  La Rosa, Thomas J.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15761)B
```

```
;  CURRENT APPLICATION NUMBER: US/09/531,113
;  CURRENT FILING DATE:  2000-03-22
;  NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.:              6.64e-28        Length:         554
Score:                  348.50          Matches:        76
Percent Similarity:     66.42%          Conservative:   13
Best Local Similarity:  56.72%          Mismatches:     26
Query Match:            14.66%          Indels:         20
DB:                     5               Gaps:           2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||              :::   ||||||   |||:::   ||||||
Db        193 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 252

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        253 AGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCTCTT 312

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||:::||||||||||
Db        313 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAAAGG 372

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              |||   |||:::   |||:::|||:::||||||||||:::|||      |||   |||:::
Db        373 AAGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTTCGG 429

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   |||   |||||||||
Db        430 AAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGCAGC--- 486

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
                                                             |||:::|||
Db        487 ------------------------------------------------------TTTAGGTTA 495

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
                 |||   ||||||   ||||||||||||||||||:::|||
Db        496 CTGGAGAT-TACCAAGCAGTGAATGTCTCCGTGAAAATTAAT 536

RESULT 14
US-09-531-113-5697
; Sequence 5697, Application US/09531113
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  La Rosa, Thomas J.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15761)B
;  CURRENT APPLICATION NUMBER: US/09/531,113
;  CURRENT FILING DATE:  2000-03-22
;  NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.:              6.64e-28        Length:         554
Score:                  348.50          Matches:        76
Percent Similarity:     66.42%          Conservative:   13
Best Local Similarity:  56.72%          Mismatches:     26
Query Match:            14.66%          Indels:         20
DB:                     6               Gaps:           2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)

Qy          1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
              ||||||||||||||||              :::   ||||||   |||:::   ||||||
Db        193 ATGTGCCGGCAAAATTCTCCGGTACACCATGCCGGTGAAGAAGAAATTGCAGCTGATGAG 252

Qy         21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
              :::|||||||||||||||||||||||||   ||||||||||||   :::   |||:::|||
Db        253 AGTCTTTTAATTTATTGCAAGCCTGTTGAACTGTACAATATTCTCTACCGCCGCGCTCTT 312

Qy         41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
                 ||||||||||||||   |||||||||   |||||||||   |||:::||||||||||
Db        313 CAAAATCCTTCTTTTCTTAGGAGATGTTTGCTTTATAAAATAAGAGCAAGGCGTAAAAGG 372

Qy         61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
              |||   |||:::   |||:::|||:::||||||||||:::|||      |||   |||:::
Db        373 AAGTTGAGAGCA---GGAATTGTGATTTTCAATTATAGGGATCACTACAACATTCTTCGG 429

Qy         81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
              ||||||||||||   ||||||   ||||||||||||||   |||   |||||||||
Db        430 AAAACTGAAGTGACCGAAGACTTTTCTTGTCCGTTTTGCTTGATGCAGTGTGGCAGC--- 486

Qy        101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
                                                             |||:::|||
Db        487 ------------------------------------------------------TTTAGGTTA 495

Qy        121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
                 |||   ||||||   ||||||||||||||||||:::|||
Db        496 CTGGAGAT-TACCAAGCAGTGAATGTCTCCGTGAAAATTAAT 536

RESULT 15
US-09-837-604B-940/c
; Sequence 940, Application US/09837604B
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  APPLICANT:  De La Pena, Robert C.
;  APPLICANT:  Bougri, Olegs
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(51892)B
;  CURRENT APPLICATION NUMBER: US/09/837,604B
;  CURRENT FILING DATE:  2001-04-18
;  PRIOR APPLICATION NUMBER: US 60/197,872
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 940
;   LENGTH: 531
;   TYPE: DNA
;   ORGANISM: Oryza sativa nipponbare
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3431-030-P1-N1-G11
US-09-837-604B-940

Alignment Scores:
Pred. No.:              9.44e-22        Length:         531
Score:                  292.00          Matches:        49
Percent Similarity:     75.31%          Conservative:   12
Best Local Similarity:  60.49%          Mismatches:     20
Query Match:            12.28%          Indels:         0
DB:                     5               Gaps:           0

US-09-890-220-2 (1-445) x US-09-837-604B-940 (1-531)

Qy        321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
              |||:::|||||||||||||||||||||||||||||:::|||||||||||||||   |||
Db        529 CATATGTGGAATTCATTTGTTCGGAAACAAAGGGTACTAGCGGATGGCCATATTCCCTGG 470
```

```
Qy        341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
              ||||||||||||||||||:::|||:::   :::||||||   :::     :::|||
Db        469 GCATGCGAAGCATTCTCGCAGTTTCATGGACAAGAACTTGTACAAAATCCAGCTCTACTA 410

Qy        361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
              ||||||||||||   |||::::::|||||||||||||||   |||:::|||:::
Db        409 TGGTGTTGGAGGTTTTTTATGGTCAAACTCTGGAACCACAGTCTACTGGATGCGCGAGCC 350

Qy        381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
              :::|||   ||||||||||||||||||         |||   |||
Db        349 ATGAATGCCTGCAACACAATTCTTGAAGGCTACCTGAACGGAAGCTCGGATCCAAAGAAA 290

Qy        401 Asn 401
              |||
Db        289 AAT 287
```

Search completed: June 20, 2003, 02:31:35
Job time : 982 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:32:03 ; Search time 42 Seconds
(without alignments)
1018.567 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INNKNNVDNKDNNSRDKVIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	629	26.5	240	2	H71435	hypothetical prote
2	423	17.8	851	2	G71435	hypothetical prote
3	322.5	13.6	623	2	D71435	hypothetical prote
4	300	12.6	582	2	E84771	fertilization-inde
5	266	11.2	295	2	A71436	hypothetical prote
6	147.5	6.2	4550	2	T18440	hypothetical prote
7	140.5	5.9	1711	2	T18429	hypothetical prote
8	133.5	5.6	3848	2	T17414	TipC protein - sli
9	132	5.6	749	2	A57276	MAD1 protein - yea
10	131.5	5.5	1245	2	D71613	GAF domain protein
11	130.5	5.5	699	2	T18426	hypothetical prote
12	128.5	5.4	1817	2	D71606	hypothetical prote
13	128.5	5.4	3844	2	T18402	asparagine/asparta
14	127.5	5.4	448	2	S05355	hypothetical prote
15	127.5	5.4	947	2	T08605	hypothetical prote
16	127.5	5.4	961	2	S67568	probable membrane
17	127.5	5.4	962	2	JC5808	G protein-coupled
18	126.5	5.3	1584	2	T18276	protein-tyrosine k
19	126.5	5.3	2150	2	S71629	sensory transducti
20	126	5.3	518	2	S23692	erythrocyte membra
21	125.5	5.3	1402	2	T17456	cell surface prote
22	124.5	5.2	686	2	A71607	Mtn3/RAG1IP-like p
23	124	5.2	1619	2	T18499	hypothetical prote
24	123.5	5.2	1858	2	T18273	1-phosphatidylinos
25	123.5	5.2	2010	2	B71616	phosphatase (acid
26	123.5	5.2	3394	2	T18501	hypothetical prote
27	122.5	5.2	2269	2	T18472	hypothetical prote
28	122	5.1	720	2	T51007	hypothetical prote
29	122	5.1	1457	2	T14577	protein kinase Yak
30	122	5.1	1714	2	E71609	Ser/Thr protein ki
31	120.5	5.1	608	2	T18437	hypothetical prote
32	120.5	5.1	1436	2	D71618	hypothetical prote
33	120	5.0	911	2	T18451	hypothetical prote
34	120	5.0	1119	2	T18491	hypothetical prote
35	120	5.0	1195	2	S38174	probable purine nu
36	120	5.0	2380	2	E71604	hypothetical prote
37	119	5.0	1094	2	S46021	probable regulator
38	118.5	5.0	1278	2	A71609	probable secreted
39	118	5.0	3724	2	T18427	hypothetical prote
40	117.5	4.9	974	1	A40213	optic lobe develop
41	117.5	4.9	1075	2	S54067	probable membrane
42	117	4.9	590	2	S66956	hypothetical prote
43	117	4.9	2175	1	S03170	homeotic protein c
44	114.5	4.8	858	1	A42239	adenylate cyclase
45	114.5	4.8	1338	2	T18287	protein-tyrosine k

ALIGNMENTS

RESULT 1
H71435
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002
C;Accession: H71435
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
. P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: H71435
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-240 <BEV>
A;Cross-references: GB:Z97342; NID:g2245031; PID:g2245039
C;Genetics:
A;Map position: 4COP9-4G3845

Query Match 26.5%; Score 629; DB 2; Length 240;
Best Local Similarity 47.6%; Pred. No. 1.1e-41;
Matches 140; Conservative 2; Mismatches 0; Indels 152; Gaps 2;

```
Qy          4 QNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRK-- 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 QNCRAKSSPEEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRKYA 110

Qy         62 ------------------------------------SRSTGMVVFNYKDCNNTLQKTEVR 85
                                                   ||||||||||||||||||:||
Db        111 FLLECSCHSDMLFILLLIWKLMNYLSLLSRYGHNEWSRSTGMVVFNYKDCNNTLQRTE-- 168

Qy         86 EDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNSFIFEEEGSDD 145

Db        169 ------------------------------------------------------------ 168

Qy        146 DKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYP 205
                                                        ||||||||
Db        169 ----------------------------------------------------GNRGLGYP 176

Qy        206 EATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEAR 259
              ||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        177 EATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTEAVVPATKTRKLSAERSEAR 230
```

RESULT 2
G71435
hypothetical protein - Arabidopsis thaliana

```
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: G71435
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: G71435
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-851 <BEV>
A;Cross-references: GB:Z97342; NID:g2245031; PID:e327517; PID:g2245038
C;Genetics:
A;Map position: 4COP9-4G3845

  Query Match             17.8%;  Score 423;  DB 2;  Length 851;
  Best Local Similarity   39.6%;  Pred. No. 6.5e-25;
  Matches  101;  Conservative    2;  Mismatches    0;  Indels  152;  Gaps    2;

Qy         43 PSFLPRCLNYKIGAKRKRK----------------------------------------SRS 64
              ||||||||||||||||||                                        |||
Db        701 PSFLPRCLNYKIGAKRKRKYAFLLECSCHSDMLFILLLIWKLMNYLSLLSRYGHNEWSRS 760

Qy         65 TGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEY 124
              ||||||||||||||||:||
Db        761 TGMVVFNYKDCNNTLQRTE---------------------------------------- 779

Qy        125 QTVNVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSL 184

Db        780 ------------------------------------------------------------ 779

Qy        185 TNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVP 244
                      |||||||||||||||||||||||||||||||||||||||||:|||||
Db        780 ------------GNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTTEAVVP 826

Qy        245 ATKTRKLSAERSEAR 259
              ||||||||||||||
Db        827 ATKTRKLSAERSEAR 841


RESULT 3
D71435
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C;Accession: D71435
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: D71435
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-623 <BEV>
A;Cross-references: GB:Z97342; NID:g2245031; PID:g2245035
C;Genetics:
A;Map position: 4COP9-4G3845
C;Superfamily: Arabidopsis thaliana 70K hypothetical protein

  Query Match             13.6%;  Score 322.5;  DB 2;  Length 623;
  Best Local Similarity   43.4%;  Pred. No. 3.1e-17;
  Matches   76;  Conservative   28;  Mismatches   38;  Indels   33;  Gaps    6;

Qy        237 LTSEAVVPATKTRKLSAERSEARSHLL-LQKRQFYHSHRVQPMALEQVMSDRDSEDEVD- 294
              ||:|| ||         |:||:| || | | |||||||   ||::||||||||||::||
Db          6 LTTEAKVP--------AKRSKATSHYLPLHKRQFYHSRTGQPLSLEQVMSDRDSENDVDK 57

Qy        295 -DDVADFEDRQMLDDFVDVNK-DEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKE 352
               || |  |: |||:  :| |:   ::|: ||||||::||::|| || |||||||| : :|
Db         58 NDDAAHLEESQMLNGSMDENEIVAERFIKLWNSFVKQQRIVADAHIPWACEAFSRLHLQE 117

Qy        353 LHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVD 407
              |    ||         | |    ::| :           ::: ||    : :|
Db        118 LRSNLSL---------DLGNTETTETAPV-----------ATEATTHGEQDGID 151


RESULT 4
E84771
fertilization-independent seed 2 protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84771
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84771
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-582 <STO>
A;Cross-references: GB:AE002093; NID:g4263788; PIDN:AAD15448.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g35670
A;Map position: 2

  Query Match             12.6%;  Score 300;  DB 2;  Length 582;
  Best Local Similarity   19.6%;  Pred. No. 1.6e-15;
  Matches  117;  Conservative   54;  Mismatches  118;  Indels  308;  Gaps   13;

Qy         13 EEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRKSRSTGMVVFNY 72
              :: :  |||:: | ||| :|     ||  |
Db         46 DDDVDVDENIIKYIKPVAVYKKLETRSKNN---------------------------- 75

Qy         73 KDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQ--TVNVS 130
                                          ||| || |||| |:|||   |:     |:||
Db         76 ----------------------------GLQLHLKSSHDAFKFEFYRAEKDHGPEVDVS 106

Qy        131 VKLNSFIFE--EEGSDDDKFEPFSLCSKPR-KRRQRGGRNNTRRLKVCFLPLDSPSLTNG 187
              || ::  |   ::   : :  | : ||| | :||||   || ::| |  : ||   |  |
Db        107 VKSDTIKFGVLKDDVGNPQLSPLTFCSKNRNQRRQRDDSNNVKKLNVLLMELDLDDLPRG 166

Qy        188 TENGITLLNDGNRGLGYPEA-----------TELAGQFEMTSNIP---------PAIAH 226
              |||  | :|| |     | |            |:||       :|         |  ||
Db        167 TENDSTHVNDDNVS-SPPRAHSSEKISDILTTTQLAIAESSEPKVPHVNDGNVSSPPRAH 225

Qy        227 SSLDAGAKVILTSEAVV---------------------------------------------- 243
              || :      :  :  |
Db        226 SSAEKNESTHVNDDDDVSSPPRAHSLEKNESTHVNEDNISSPPKAHSSKKNESTHMNDED 285

Qy        244 ----PATKTRKLSA-----------ERSEARSHLLLQKRQF----YHSHRVQPMALE--- 281
                  | |:: | ::           | || :   : :::|     | :   ||   |
Db        286 VSFPPRTRSSKETSDILTTTQPAIVEPSEPKVRRVSRRKQLYAKRYKARETQPAIAESSE 345

Qy        282 ---------------------------------------------------------- 281

Db        346 PKVLHVNDENVSSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDENVSSTPRAHSSK 405

Qy        282 QVMSDRDSEDEV----------------------------------DDDVAD------- 299
              :  | | : | |                                  ||:|:
Db        406 KNKSTRKNVDNVPSPPKTRSSKKTSDILTTTQPTIAESSEPKVRHVNDDNVSSTPRAHSS 465

Qy        300 ----------------------------------------------------FEDRQ 304
```

```
Db        466 KKNKSTRKNDDNIPSPPKTRSSKKTSNILATTQPAKAEPSEPKVTRRLKGRQFYHSQTMQ 525

Qy        305 MLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFW 361
Db        526 RLERLVGVSKEEKRYMYLWNIFVRKQRVIADGHVPWACEEFAKLHKEEMKNSSSFDW 582


RESULT 5
A71436
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: A71436; B71436
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: A71436
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-295 <BEV>
A;Cross-references: GB:Z97342; NID:g2245031; PID:e327018; PID:g2245041
A;Accession: B71436
A;Molecule type: DNA
A;Residues: 1-295 <BE2>
A;Cross-references: GB:Z97342; NID:g2245031; PID:e327518; PID:g2245040
C;Genetics:
A;Map position: 4COP9-4G3845

  Query Match             11.2%;  Score 266;  DB 2;  Length 295;
  Best Local Similarity   53.4%;  Pred. No. 3e-13;
  Matches   70;  Conservative    3;  Mismatches   30;  Indels   28;  Gaps    4;

Qy        278 MALEQVMSDRDSEDEVDDDVADFEDRQ------------------MLDDFVDVNKDEKQ 318
Db          1 MALEQVMSDRDSEDEVDDDVADFEDRQVFHDFFLRSLSRQQKMVYDMLDDFVDVNKDEKQ 60

Qy        319 FMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDS 378
Db         61 FMHLWNSFVRKQR---NKHILLPIFAFIAIPLAALH--LSL----TLTSFRLKNHVFRLE 111

Qy        379 ATINNCNTILE 389
Db        112 ALANVVHTRFE 122


RESULT 6
T18440
hypothetical protein C0425w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T18440
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18440
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4550 <LAW>
A;Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325396; PIDN:CAB11121.1
C;Genetics:
A;Map position: 3
A;Note: C0425w

  Query Match              6.2%;  Score 147.5;  DB 2;  Length 4550;
  Best Local Similarity   19.8%;  Pred. No. 0.018;
  Matches   81;  Conservative   64;  Mismatches  150;  Indels  115;  Gaps   18;

Qy         66 GMVVFNYKDCNN------TLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFK 119
Db       1820 GLYTYNLKNNNNNNVYMKTIKLT------------LCCT------HSNEHDTKAVMNIK 1860

Qy        120 L----FEEYQTVNVSV--KLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLK 173
Db       1861 LNTLNNKKYNNINCIFYHYYNSFLHSLFYSD----EYIELIRK-EERKEDIEKEKIKRGK 1915

Qy        174 VCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGA 233
Db       1916 YCHQEDRDDNMNNNDNNDNNNNNDNNNNNDNNNNND--------NNNNDNNNNDNNNSGE 1967

Qy        234 KVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEV 293
Db       1968 DQLINED-----NKKRKINNLSTHEHTHNM-------SDHNKEP----KKKKNKSNETQI 2011

Qy        294 DDDV-------ADFEDRQMLDDFVD---------VNKDEKQFMHLWNSFVRKQRVIADGH 337
Db       2012 NDDTNKKETHKNEIDHKEQKDTNCDNGIEEKKDQVSKSNNIISKVFHFFERKNNNI---- 2067

Qy        338 ISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDT 397
Db       2068 ------------KKDEHK------------KEYNE---RNDKMSECDMLLQNKRDNDHN 2099

Qy        398 TTTNNNNSVD------RPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRD 441
Db       2100 NNNNNNNDDDNNNNNNNDDDNNNNNNNNDDHNNN-NNNNDDDNNNNNNND 2148


RESULT 7
T18429
hypothetical protein C0345w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T18429
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18429
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1711 <LAW>
A;Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325381; PIDN:CAB11106.1
C;Genetics:
A;Note: PFC0345w

  Query Match              5.9%;  Score 140.5;  DB 2;  Length 1711;
  Best Local Similarity   18.5%;  Pred. No. 0.018;
  Matches  105;  Conservative   95;  Mismatches  204;  Indels  165;  Gaps   21;

Qy          4 QNCRAKSSPEEVISTD-----ENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLN--YKIGA 56
Db       1054 QSNNTNSKDEHNISSDKSKKEDTLNLSRKSSYEYNNKILQSTSNKS-----LNGAYENNL 1108

Qy         57 KRKRKSRSTGMVVFNYKDCNNTLQKTEVREDCSCPFCSMLCGSFK-----GLQFHLNSS- 110
Db       1109 FSGKKKKNKGTVLKDIEHINDIQDKYPEDLNINCVNKYVIENEEKHLLPLELEYNLVSSD 1168

Qy        111 ----------------------HDLFEFEFK---LFEEYQTVNVS--------VKLNSFI 137
Db       1169 EKFGLNKIKNDNNIIYMKHQNYHNLYDDNQKKHILFDTNKNVSIQRNNNINSVIKTNHYE 1228

Qy        138 FEEEGSDDDKFEPFSLCSKPRK--------------------RRQRGG-------RNNTR 170
Db       1229 VEKNNKDQRNYDNFT-CDKKKKIYYNIINSDKDIYHNNIIYTKNEKEGIGNIHLNRNDKD 1287

Qy        171 RLKVCFLPLDSP------------SLTNGTENGITLLNDGNRGLGYPEATELAGQ-FEMT 217
Db       1288 ITNFELLKLDGVKEFLDTFKDSYIDCHNKKEN---ILNMTNKNKEDHQIIDVADKIFNET 1344

Qy        218 SNIP-------------------PAIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEA 258
```

```
Db       1345 NMITMDNNKIYDDKNVHEKKCTHNDVIHHNMDI------------------LSTSIKNN 1385

Qy        259 RSHLLLQKRQFYHSHRVQPMALEQ--VMSDRDSEDEVDDDVADFEDRQML--------DD 308

Db       1386 EENLFIDTYQ--KQNRIGDIYMNRINILQEDDDDDNHNNNHNNNNNNNKLILFEYTKNDQ 1443

Qy        309 FVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLI 368

Db       1444 MLHNNKNNLEGTEEFSDFIEKKNKI---KIKNKNESYHKIDESLLSNEKN-----NKVSL 1495

Qy        369 KLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINN 428

Db       1496 LLINNNKDSSSVDNNKNNNNKNNNNKNNNNENNNKNNKNNNNDSFSKDNNLINNDNNNNN 1555

Qy        429 K------------NNVDNKDNNSRDKVIK 445

Db       1556 NNNDSFSKDNNLINNDNNNNNNNNNKVIK 1584
```

RESULT 8
T17414
TipC protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17414
R;Stege, J.T.; Laub, M.T.; Loomis, W.F.
submitted to the EMBL Data Library, July 1998
A;Description: Interaction of tip genes in early Dictyostelium discoideum development.
A;Reference number: Z18774
A;Accession: T17414
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3848 <STE>
A;Cross-references: EMBL:AF079445; NID:g3420744; PID:g3420745; PIDN:AAC31916.1
A;Experimental source: strain AX4
C;Genetics:
A;Gene: tipC
A;Introns: 72/3
C;Function:
A;Description: required for tip formation

  Query Match 5.6%; Score 133.5; DB 2; Length 3848;
  Best Local Similarity 19.7%; Pred. No. 0.18;
  Matches 75; Conservative 47; Mismatches 103; Indels 155; Gaps 12;

```
Qy          76 NNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNS 135

Db        2040 NNSNNQLPVKEDSQ------------LQKYLSISLEKGEF---ILNDDHKISSPIKLLS 2083

Qy         136 FIFEEEGSDDDKFEPPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTENGITLL 195

Db        2084 I-----GVDGLKSNIFSF---PQK-----------------------------NQIALS 2105

Qy         196 NDGNRGLGY--------PEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTSEAVVPATK 247

Db        2106 LDANMKAGYFNKNIGIWEPLIENWG-FSFTSN-------NSIEGGWMVNFNSKIPLYINI 2157

Qy         248 TR-----KLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFED 302

Db        2158 TKIFIDTSISTYQIWADDYYSQQKKD-----------KKNKSNDYDNDEIIEDTV----- 2201

Qy         303 RQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSLFWC 362

Db        2202 EQVNNDLIKPADDQK---------------------------------------------- 2216

Qy         363 WRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDH 422

Db        2217 ------------------NNNNNKNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2257

Qy         423 PNDINNKNNVDNKDNNSRDK 442

Db        2258 NNNNNNNNKNNNDNNNNIK 2277
```

RESULT 9
A57276
MAD1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein G3191; protein YGL086w
C;Species: Saccharomyces cerevisiae
C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 21-Jul-2000
C;Accession: A57276; S64093; S48246
R;Hardwick, K.G.; Murray, A.W.
J. Cell Biol. 131, 709-720, 1995
A;Title: Madlp, a phosphoprotein component of the spindle assembly checkpoint in budd
A;Reference number: A57276; MUID:96042315; PMID:7593191
A;Accession: A57276
A;Molecule type: DNA
A;Residues: 1-749 <HAR>
A;Cross-references: GB:U14632; NID:g551091; PIDN:AAA91620.1; PID:g551092
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64093
A;Molecule type: DNA
A;Residues: 1-749 <RIE>
A;Cross-references: EMBL:Z72608; NID:g1322610; PID:e243297; PID:g1322611; MIPS:YGL086
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:MAD1
A;Cross-references: SGD:S0003054; MIPS:YGL086w
A;Map position: 7L
C;Function:
A;Description: component of spindle assembly checkpoint which prevents cells from ini
C;Keywords: cell division control; nucleus; phosphoprotein

  Query Match 5.6%; Score 132; DB 2; Length 749;
  Best Local Similarity 21.3%; Pred. No. 0.029;
  Matches 87; Conservative 57; Mismatches 143; Indels 122; Gaps 18;

```
Qy          89 SCPFCSMLCGS-------------FKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLNS 135

Db          40 SSPFLESPGGSPDVGSTNGQSNRQIQALQFKLNTLQNEYEIE-KLQLQKQTNILEKKYKA 98

Qy         136 FIFE-EEGSDDDKF-----EPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTNGTE 189

Db          99 TIDELEKALNDTKYLYESNDKLEQELKSLKERSANSMNDKDK---CIEELRT-------- 147

Qy         190 NGITLLNDGNRGLGYPEATELAGQFE-MTSNIPPAIAHSSLDAGAKVILTSEAVVPATKT 248

Db         148 ---TLQNKD------LEMETLRQQYDSKLSKVTNQCDHFKLEAES--------------- 183

Qy         249 RKLSAERSEARSHLLLQ------KRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVADFED 302

Db         184 -----------SHSLLMKYEKEIKRQSVDIKDLQ----HQVM-------EKDDELSSVKA 221

Qy         303 RQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIAD----------GHISWACEAFSRFYEKE 352

Db         222 SKMINSHPNYSTEE------FNELTEMNKMIQDQVQYTKELELANMQQANEL------KK 269

Qy         353 LHRYSSLFWCWRLFLIKLWN-----HGL---VDSATINNCN--------TILENCRNSSD 396

Db         270 LKQSQDTSTFWKLENEKLQNKLSQLHVLESQYENLQLENIDLKSKLTKWEIYNDSDDDDD 329

Qy         397 TTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445

Db         330 NNVNNNDNNNNNKNDNNNDNNNDTSNNNNINNNNRTKNNIRNNPEEIIR 378
```

RESULT 10
D71613
GAF domain protein (cyclic nt signal transduct.) PFB0510w - malaria parasite (Plasmod
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C;Accession: D71613
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

```
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: D71613
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1245 <GAR>
A;Cross-references: GB:AE001399; GB:AE001362; NID:g3845200; PIDN:AAC71891.1; PID:g38452
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0510w

  Query Match             5.5%;  Score 131.5;  DB 2;  Length 1245;
  Best Local Similarity   18.4%;  Pred. No. 0.062;
  Matches   89;  Conservative   78;  Mismatches  161;  Indels  155;  Gaps   21;

Qy         57 KRKRKSRSTGMVVFNY----KDCNNTLQKTEVREDCSCPFCSMLCGSFKGLQFHLNSSHD 112
              |:|:| :   :: :|     || :   : :   :                  ||   |
Db        530 KKKKKRQKRNILPYNILKMEKDDDKIFDEGDENRESK---------------ELNEDED 573

Qy        113 LFEFEFKLFEEYQTV--NVSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTR 170
              | | :  || : |    ::| |    ::   ||||::             :   :: :
Db        574 EEEDEEEDEEEDEEVVYEENIKENK---DDIDYDDDKYD------------KNYEHDEK 617

Qy        171 RLKVCFLPLDSPSLTNG---TENGITLLNDGNRGLGYPEATELAG-----QFEMTSNIPP 222
              :  : :          ||    ||   |: |      | |   :       |:|  :||
Db        618 KENIFY-------KQNGDDHDENNYELIQD----YYYVEKNNIYNKNDTIQYEQNNNI-- 664

Qy        223 AIAHSSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQ-KRQFYHSHRVQPMALE 281
                      |   ::    | |       :  : |     | |    :  : | |
Db        665 ----------YKTKFSNNIYTP--KFDNYMDYKGNIKEHTANNIKHQLLLFNIDKTMKDE 712

Qy        282 QVMSDRDSEDEVDDDVADF-----EDRQMLDDFVDVNK----DEKQFMHLWNSFVRKQRV 332
              : ||: ::|:  ||: :       |  ::      ::|    |: |    :   : | :
Db        713 ILTSDKINKDKSSDDIINSNKTYNEKNKLYTSTTGIHKNNMGDKLQDDTFYFKLLNKNKY 772

Qy        333 IADG--------HISWACEAFSRFYEKELHRYSSLF---WCWRLFL------IKLWN-HG 374
              :          :| :  |:  : ::|       ||    | |:|:      :|  | :
Db        773 YMNNVNGNLDIYYILYNIESVEQLFKKIKENNFLLFNMKNCIRMFILFRDLYVKEENGYN 832

Qy        375 LVDSATINNCNTI--LENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIV------------ 420
              :: |   ||  ||     | :|:|::  ||||: :  :::| |||| |
Db        833 IIKSQ--NNTTTITTTTNSNDSNDSSDNNNNNNNNNNNNNNNYNNNNSVIFSTNEKIYDML 890

Qy        421 ----------------------------------------DHPNDINNKNNVDNKDNNSR 440
                                                       |: |: || ||:|| :||:
Db        891 NRDNIYKKVKKEIFEGDSIIKTMENKPNLTNKNYMNNDNIDNNNNNNNNNNIDNNNNNNG 950

Qy        441 DKV 443
              | :
Db        951 DNI 953


RESULT 11
T18426
hypothetical protein C0330w - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T18426
R;Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z18935
A;Accession: T18426
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-699 <LAW>
A;Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325377; PIDN:CAB11102.1
C;Genetics:
A;Note: C0330w

  Query Match             5.5%;  Score 130.5;  DB 2;  Length 699;
  Best Local Similarity   25.6%;  Pred. No. 0.035;
  Matches   43;  Conservative   26;  Mismatches   62;  Indels   37;  Gaps    6;

Qy        289 SEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRF 348
              |||| :  :  | |:   || :    |:    :   :   |  :|         |
Db        112 SEDEKESSIDIFHDKINDDDLMIYEDDDDAKKYFGKTTFNKPNIIQKD------ENSEDD 165

Qy        349 YEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNN------ 402
                : :|:  :                    |    |  | : :|  |:|    |||
Db        166 QDDDLNDQN------------------DDPINNPTNNLTDNPTNNSTDNPTNNSIDNPT 206

Qy        403 NNSVDRPSDSNTNN--NNIVDHP--NDINN--KNNVDNKDNNSRDKVI 444
              |||:| |:::: :|  || :|:|  | |:|   |::||  ||| |  |
Db        207 NNSIDNPTNNSIDNPTNNSIDNPTNNSIDNPTNNSIDNPTNNSTDNPI 254


RESULT 12
D71606
hypothetical protein PFB0800c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C;Accession: D71606
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: D71606
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1817 <GAR>
A;Cross-references: GB:AE001418; GB:AE001362; NID:g3845275; PIDN:AAC71949.1; PID:g384
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0800c

  Query Match             5.4%;  Score 128.5;  DB 2;  Length 1817;
  Best Local Similarity   17.5%;  Pred. No. 0.17;
  Matches   83;  Conservative   79;  Mismatches  160;  Indels  151;  Gaps   18;

Qy         25 YCKPVRLYNIFHLRSLGNPSFLPRCLNYKIGAK---------RKRKSRSTG---MVVFNY 72
              | | ::| :|| | : |   ||    | ||  :          :| |:  |    ::  |
Db        382 YMKHIKLEHIFELFTEGCYLFL--LPNEKIKSNNIYTPNIFLKKLKNYITSNDFVIHIND 439

Qy         73 KDCNNTLQKTEVREDCSCPFCSMLCGS-------FKGLQFHLN-----SSHDLFEFEFK- 119
              :  |  ::           | |  | :       |  |   |:     :::|:    :|
Db        440 RTLNRYIK-----------FLSYHCSNNIYVHILFNDLYITLHKKIFINNYDMLIKHYKS 488

Qy        120 -------LFEEYQTVNV--SVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTR 170
                    |  : : :|   :: ||::              :|| :|     :|    |
Db        489 TADHILYLTNKNENLNYLNTILLNNY-------------YSLYNKKENNEKRQSLEN-- 532

Qy        171 RLKVCFLPLDSPSLTNGTENGITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLD 230
               |||   |       : |   :   || |:   : :       :|   |
Db        533 -LKVKIFPSHLNYKQHTTNKNV---NDPNQQHKHDKDDSYDNTYEQMKN---------- 577

Qy        231 AGAKVILTSEAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSE 290
                      : |                          ::  :| :|    : : | :  :
Db        578 -------NKNKIYP---------------------NEYITTHILQNNYEQNLYSFQKKD 608

Qy        291 DEVDDDVADFEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYE 350
              |   ::: |   |: : ::   |:   |   :   :::::              :
Db        609 DTNINNIFDLHKREQIYEYEKENESSDIFRDSYKRKIKEEK----------------KK 651

Qy        351 KELHRY-----------SSLFWCWRLFLIKLWNHGLVDSATINNC--NTI----LENCR 392
              | :::|            |:          |:   |    |:.  | |    |||  :
Db        652 KNIYKYEDHPLNKEKKKKKKFFYI---------NYEKGDDKNDNDLYYNNIYSKNLENIQ 702

Qy        393 NSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNSRDKVIK 445
              | :     ||||: :  :::| ||||  :: |: || || :| :||: :   |
Db        703 NKNYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNFDK 755


RESULT 13
```

```
T18402
asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T18402
R;Barale, J.C.; Candelle, D.; Attal-Bonnefoy, G.; Dehoux, P.; Bonnefoy, S.; Ridley, R.;
Infect. Immun. 65, 3003-3010, 1997
A;Title: Plasmodium falciparum AARP1, a giant protein containing repeated motifs rich in
A;Reference number: Z18929; MUID:97378065; PMID:9234746
A;Accession: T18402
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3844 <BAR>
A;Cross-references: EMBL:Y08926; NID:e1154302; PID:e311435; PIDN:CAA70130.1
C;Genetics:
A;Gene: aarp1

  Query Match             5.4%;  Score 128.5;  DB 2;  Length 3844;
  Best Local Similarity   19.1%;  Pred. No. 0.45;
  Matches   85;  Conservative   49;  Mismatches  141;  Indels  169;  Gaps   14;

Qy         30 RLYNIFHLRSLGNPSFLPRCLNYKIGAKRKRKSRSTGMVVFN---------------YKD 74
              :| |:|  :   :|||   |  | | |      :  ::  |                 :||
Db        992 QLENVFPKQFRHHPSFYE--LVNKYGYKTHMPRSNVNLIKLNKNSWYLYDAFWPLSAFKD 1049

Qy         75 CNNTLQKTEVREDCSCPFCS------MLCGSFKGLQFHLNSSHDLFEFEFKLFEEYQTVN 128
                 :|     |: |  ||       |  ||: :|       :|  : ||      :
Db       1050 FQTANEKCIKEENSSYIGCSREEDKEYLNFSFRKVQ-------NLLFYTFK-----NSCL 1097

Qy        129 VSVKLNSFIFEEEGSDDDKFEPFSLCSKPRKRRQRGGRNNTRRLKVCFLPLDSPSLTN-- 186
              ||: |  ||  :|         |    : :|: :       : :|       |  || |
Db       1098 VSILLIVFIINDE---------FQRKIEKKKKEELEKEEMKKHMKADVTGEDD-SLLNQK 1147

Qy        187 -GTEN------GITLLNDGNRGLGYPEATELAGQFEMTSNIPPAIAHSSLDAGAKVILTS 239
               || |      |: :  : ||  |            |: |                 ::|
Db       1148 GGTTNNNTCDVGVVINEEANRSSG-----------MSQN---------------VVTY 1179

Qy        240 EAVVPATKTRKLSAERSEARSHLLLQKRQFYHSHRVQPMALEQVMSDRDSEDEVDDDVAD 299
              |: :      |  ||:                       | | | ::: :||  |
Db       1180 ESGINMNNDDVLGEERTNNDD-------------------INYMDDNDDDNDDNDDDND 1219

Qy        300 FEDRQMLDDFVDVNKDEKQFMHLWNSFVRKQRVIADGHISWACEAFSRFYEKELHRYSSL 359
                |    ||  | | |:
Db       1220 DNDDDNDDDNNDDNNDD------------------------------------------ 1236

Qy        360 FWCWRLFLIKLWNHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNI 419
                                 || |   ::  |  |    :||:  :   |:| ::||
Db       1237 ----------------------NNDNDNNDDDNNDDDNNDDDNNDDDNNDDDNNDDDNND 1274

Qy        420 VDHPNDINN--KNNVDNKDNNSRD 441
               |: :| ||   || |: |:|  |
Db       1275 DDNNDDDNNDDDNNDDDNDDNDED 1298


RESULT 14
S05355
hypothetical protein (clone AAC11) - slime mold (Dictyostelium discoideum) (fragment)
C;Species: Dictyostelium discoideum
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 29-Oct-1999
C;Accession: S05355
R;Shaw, D.R.; Richter, H.; Giorda, R.; Ohmachi, T.; Ennis, H.L.
Mol. Gen. Genet. 218, 453-459, 1989
A;Title: Nucleotide sequences of Dictyostelium discoideum developmentally regulated cDNA
A;Reference number: S05355; MUID:90066348; PMID:2511421
A;Accession: S05355
A;Molecule type: mRNA
A;Residues: 1-448 <SHA>
A;Cross-references: EMBL:X16522; NID:g7173; PIDN:CAA34529.1; PID:g7174

  Query Match             5.4%;  Score 127.5;  DB 2;  Length 448;
  Best Local Similarity   35.7%;  Pred. No. 0.034;
  Matches   25;  Conservative   20;  Mismatches   24;  Indels    1;  Gaps    1;

Qy        372 NHGLVDSATINNCNTILENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNN 431
              |:   ::: ||| |     |  |::: :: ||||: :  :::||||||   : |  || ||
Db        273 NNNNNNNSNINNINNNNNNNSNNNNNSSNNNNNN-NNSTNNNTNNNNNNTNNNTNNNNNN 331

Qy        432 VDNKDNNSRD 441
              ::| :||: :
Db        332 INNNNNNTNN 341


RESULT 15
T08605
hypothetical protein HelE - slime mold (Dictyostelium discoideum) (fragment)
C;Species: Dictyostelium discoideum
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C;Accession: T08605
R;Loomis, W.F.; Iranfar, N.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z16451
A;Accession: T08605
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-947 <LOO>
A;Cross-references: EMBL:AF019981; NID:g2425112; PID:g2425113
A;Experimental source: strain AX4
C;Genetics:
A;Gene: helE

  Query Match             5.4%;  Score 127.5;  DB 2;  Length 947;
  Best Local Similarity   23.4%;  Pred. No. 0.089;
  Matches   40;  Conservative   37;  Mismatches   87;  Indels    7;  Gaps    3;

Qy        273 HRVQPMALEQVMSDRDSEDEVDDDVADFEDRQMLDDFVDVNKDEKQFMHL----WNSFVR 328
              |  : :  |: :|  | |: :||    | | :     :| :|||::  :      |  ::
Db        116 HEARRLEFEKYISIDDDEEIIDD--IPFVDSENCRITIDFSKDEREDDYYEDNEINRIIK 173

Qy        329 KQRVIADGHISWACEAFSRFYEKELHRYSSLFWCWRLFLIKLWNHGLVDSATINNCNTIL 388
              :|          | :  :  :  ::    :|     |    :|    |    |    :
Db        174 RQTTSTSTTTSISTTSTPQKQQQNTPNHSP-SPSQRSKTTRLQQTPLTTPQTTPQTSPSS 232

Qy        389 ENCRNSSDTTTTNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVDNKDNNS 439
              :|  |:::    ||||: :  :::| ||||  :: |: || || :|  ||:
Db        233 QNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSNNN 283


Search completed: June 13, 2003, 15:38:03
Job time : 46 secs
```